On the Optimal Power Allocation for Broadcasting in Dense Cooperative Networks

Birsen Sirkeci Mergen and Anna Scaglione
Cornell University, Ithaca, NY 14853
Email: bs233@cornell.edu, anna@ece.cornell.edu

Abstract—We study the optimal power allocation problem for cooperative broadcast in dense large-scale networks. In our set-up, a single source initiates the transmission and the nodes that receive the message with sufficient signal-to-noise ratio (SNR) retransmit. The goal is to design the transmission order (schedule) and the transmission powers of the relays so that the message reaches to the entire network with the minimum possible power expenditure. In general, finding the best scheduling in cooperative broadcast is known to be NP-complete. In this paper, we show that the optimal scheduling is easily resolved for certain network topologies and channel models. Furthermore, we approximate dense large-scale networks with a continuum of nodes and provide optimal power density for the continuum network under pathloss attenuation.

I. INTRODUCTION

In large-scale networks, nodes are constrained in their size and battery power, and hence, it is crucial to design energy efficient and low complexity transmission schemes. Cooperative transmission enhances the energy efficiency either by providing diversity or by increasing the received SNR [1]–[12]. The essence of cooperative broadcast was first introduced in [4]–[6] as an energy efficient alternative to previous broadcasting schemes [13], [14]. The main idea is to utilize multiple replicas of the same message from different transmitters.

In this paper, we study the power efficiency of cooperative broadcasting in dense networks. In the optimal cooperative broadcasting (OCB), all the nodes utilize all the previous receptions [4]–[6]. In addition, the nodes transmit based on predetermined schedule and power allocation policy such that total power consumption of the network is minimized. In [4]–[6], it was shown that for a given transmission schedule, the optimal power allocation can be formulated as a constrained optimization problem which can be solved in polynomial time by utilizing linear programming tools. On the other hand, the authors showed also that finding the optimal scheduling that leads to the minimum total power consumption is an NP-complete problem and thus, it is not computationally tractable. Both works proposed heuristic methods to determine the optimal schedule. In general, OCB requires a central control unit.

First, we study specific network topologies and channel models for which we are able to show that the optimal scheduling is resolved in polynomial time. Then, we extend the analysis to dense networks. In particular, for dense large-scale networks, we approximate the optimal schedule with the schedule that allows the nodes to transmit in the order of their distances from the source node. This approximation becomes exact in the asymptote as the node density increases, which we will refer to as the continuum network. Under the continuum model, we are able to show that the optimal power density is given by the solution of a Volterra equation with parameters that depend on the network topology and the channel gains. In addition, for specific path loss models and topologies we are able to find closed form expressions for the optimal power density \( p(r) \).

The organization of the paper is as follows. In Section II, we describe the system model. In Section III, we provide the general formulation for the optimal power allocation of the cooperative broadcast scheme and we provide conditions under which the general formulation is simplified. In Section V, we derive the optimal power density under the continuum approximation of dense networks. The paper conclusion follows in Section VI.

II. SYSTEM MODEL

We consider a network formed by a single source and \( N \) relays, which are distributed randomly and uniformly in a given region. The source node initiates the transmission session. The relays which have received source message reliably at the \( k \)th time-slot are allowed to retransmit the message in the \((k+1)\)th time slot. Nodes are half-duplex, i.e. can not receive and transmit at the same time. Note that relays do not transmit the same packet more than once.

Definition The relays that are allowed to retransmit at the \( k \)th time instant are called level-\( k \) nodes. We will denote the set of level-\( k \) nodes by \( S_k \).

In this paper, we only consider a single-shot communication. We assume the nodes are stationary. We also assume that

\(^1\)This manuscript is published in ISIT 2006 without the appendix.
appropriate channel coding is used so that the decoding and retransmissions are correct as long as the received cumulative SNR is above the threshold. A training preamble in the message helps nodes to detect the packet’s presence, estimate the received power and synchronize the retransmissions in a level. Next, we describe the reception model.

A. Reception Model

Let the $i$th node transmit with power $P_i$ and let $H_{ij}$ be the deterministic link power gain. Define the received power due to transmission of nodes in $L$ as

$$Power_j = \sum_{i \in L} P_i H_{ij}. \quad (1)$$

This model is valid if the relays transmit in orthogonal channels, as in TDMA, FDMA or CDMA, or if the relays use orthogonal space-time codes as considered in [3] and the receiver is an optimum maximal-ratio-combiner (MRC) receiver [15]. In case of orthogonal channels, a large bandwidth is required, i.e., the network should operate in the wideband regime [4]. Furthermore, in this case, a centralized scheduler, which assigns orthogonal channels to the nodes, is required. Note that when space-time codes are used, the scheduling problem can be resolved via randomization [16].

If the simultaneous transmissions are not in orthogonal dimensions, the cumulative power of transmitted packets depends on the relative delays and phases of individual overlapping signals. If we assume that the transmitted signal is narrowband, then (1) represents the average power. In this scenario, the scheduling is decentralized; however, the synchronization assumptions are stricter.

In our analysis, we assume that each node can combine the packets that were received in different time intervals though maximal ratio combining. Note that we assume the channel coefficients are time-invariant and are perfectly estimated (channel state information is available only at the receiver, i.e., the transmitter does not know the channel state information).

III. Power Allocation for OCB: Problem Formulation

In the optimal cooperative broadcast scheme, the source node initiates the transmission by sending a packet. Each node accumulates signal powers from all the nodes that transmitted previously (see Section II-A). The nodes that have received sufficient signal power $\tau$, i.e., $Power \geq \tau$, are allowed to retransmit according to a given schedule. The $Power$ is defined in (1) and $\tau$ depends on the performance metric (e.g. outage capacity, bit error rate). We assume that the noise is of unit power; hence, $\tau$ will also be called the SNR threshold.

Let $I = \{1, \ldots, N + 1\}$ denote the set of node indices, where the source node is denoted by 1. The transmission schedule will be represented by a mapping $S$, i.e., $S: \{1, \ldots, N + 1\} \rightarrow \{1, \ldots, N + 1\}$. Let $P_i$ be the transmission power of the $i$th node. Let $H$ be the channel matrix such that its $(i,j)$th entry denotes the channel gain from the $j$th node to the $i$th node. Our aim is to find the best schedule $S$ and optimal power allocation $\{P_i, \forall i \in I\}$ for a given network with channel matrix $H$ (whose entries are the link power gains) and decoding threshold $\tau$ such that $\sum_i P_i$ is minimized under the described cooperative broadcast scheme. Note that the above mapping excludes the scenarios where nodes are allowed to transmit together; however, such scenarios can be easily mapped into the above formulation.

We will associate a given schedule $S$ with a permutation matrix $S \triangleq [e_{S(1)} e_{S(2)} \ldots e_{S(N+1)}]$, where $e_i$ is the $i$th column of the identity matrix. The optimal $S$ and power allocation vector $p \triangleq [P_1, P_2, \ldots, P_{N+1}]$ are the solutions of the following linear optimization problem (see also [4]–[6]):

$$\min 1^T p \quad \text{subject to} \quad L(SHST)p \geq \tau b, \quad p \geq 0, \quad (2)$$

where $b = [1; 0]$, $1$ denotes a vector of all 1’s, 0 denotes a vector of all 0’s and $[x; y]$ denotes column concentration of vectors $x$ and $y$. The operator $L$ models the causality in the system and it is defined as follows: Let $a_{i,j}$ be the $(i,j)$th element of $N \times N$ matrix $A$ and $l_{i,j}$ be the $(i,j)$th element of $L(A)$, then for $1 \leq i,j \leq N$,

$$l_{i,j} = \begin{cases} a_{i+1,j} & \text{if } i \geq j, \ i < N, \\ 1 & \text{if } i = N, \ j = N, \\ 0 & \text{otherwise.} \end{cases}$$

The last scheduled node does not need to transmit, i.e., the optimal $P_{S(N+1)} = 0$. This is enforced in the above formulation via the definitions of $b$ and $L$. In the following, we will denote the set of all possible permutation matrices with $P_S$, where $|P_S| = N!$ (the transmission is initiated by the source node, hence out of $(N + 1)!$ possibilities $N!$ of them represent valid schedules).

For a given permutation matrix $S$, the constrained optimization problem (2) can be solved in polynomial time as a function of the number of relays $N$ by utilizing efficient linear programming algorithms. However, finding the optimal scheduling, i.e., finding the best permutation matrix $S$ out of $N!$ possibilities was shown to be an NP-complete problem [4]–[6]. Hence, the problem is intractable in general. In the next section, we provide conditions under which the best scheduling is easily determined and problem (2) can be solved in polynomial time. Note that in general, OCB requires a central controller that knows all of the channel gains.

IV. Further Results on the Best Scheduling for Cooperative Broadcast

In Fig. 1(a), we present a linear network where the source node is located at the edge. For this network, under the effect of pathloss attenuation, for example $\ell(r) = 1/r^2$, the optimal schedule is such that the nodes transmit in the order of their distances from the source node [5], [6]. In the following, this will be referred as the trivial scheduling. We ask the question if there exist other networks where trivial scheduling is optimal. In Lemma 1, we provide sufficient conditions on the channel matrix $H$ so that the overall complexity of the problem (2) is decreased considerably.
Lemma 1: Consider the optimal power allocation and scheduling problem in (2). Assume that \( \mathbf{H} \) has non-negative values. Let \( \mathcal{P}_S \) denote the set of all possible permutation matrices.

a) If there exits a permutation matrix \( \hat{\mathbf{S}} \in \mathcal{P}_S \) such that \( \mathcal{L}(\hat{\mathbf{S}}\mathbf{H}^T\mathbf{S}^T) \) is column-ordered \(^2\), then \( \hat{\mathbf{S}} \) corresponds to the optimal schedule.

b) If there exits a permutation matrix \( \hat{\mathbf{S}} \in \mathcal{P}_S \) such that \( \mathcal{L}(\hat{\mathbf{S}}\mathbf{H}^T\mathbf{S}^T) \) is both column- and row-ordered \(^3\), then the optimal power allocation can be obtained by solving

\[
\mathcal{L}(\hat{\mathbf{S}}\mathbf{H}^T\mathbf{S}^T)\hat{\mathbf{S}}\mathbf{p} = \tau\mathbf{b}. \tag{3}
\]

Proof: Appendix I.

Lemma 1a) provides a sufficient condition on the channel matrix \( \mathbf{H} \) to determine the optimal schedule. The validity of the sufficient condition can be determined by ordering the first column of \( \mathbf{H} \) and by checking if the rest of the columns are ordered. This sorting and comparing algorithm has complexity \( O(N^2) \). Using Lemma 1a), we can determine that the trivial scheduling is the best schedule for both the network topologies in Fig. 1 under a pathloss attenuation model. Lemma 1b) provides a sufficient condition so that the optimal power allocation problem has the same complexity as inverting a lower triangular matrix, which is \( O(N^2) \).

For the linear networks and linear-like configurations (among the two dimensional networks, see Fig. 1), the trivial scheduling is optimal under an appropriate pathloss attenuation model. Furthermore, these observations provide us the intuition and background that will be utilized for dense large-scale networks in the next section. The interesting fact is that trivial scheduling tends to be satisfied in dense networks.

V. OPTIMUM COOPERATIVE BROADCAST IN DENSE NETWORKS

In this section, we consider the problem of optimal power allocation for dense networks under cooperative broadcasting. Suppose that \( N \) nodes are uniformly and randomly distributed within \( S = \{(x, y): x^2 + y^2 \leq R^2\} \) and the source node is located at the origin (see Fig. 2). In our analysis, we consider only the effect of pathloss attenuation on the channel gain. Let \( \ell(\cdot) \) denote the pathloss attenuation function. For a transmitter that is located at \((x, y)\) and a receiver that is located at \((x', y')\), we assume that \( \ell(\cdot) \) is

a1) a function of the distance \( d := \sqrt{(x - x')^2 + (y - y')^2} \) between the transmitter and the receiver;

a2) continuous, non-negative and decreasing in \( d \);

a3) circularly symmetric. We will use the notations \( \ell(x - x', y - y') \) and \( \ell(d) \), interchangeably.

Consider the network topologies in Fig. 3. Under the pathloss attenuation that satisfies assumptions a1)-a3), the trivial scheduling is optimal in these scenarios and the best schedule assigns the relays that are positioned at the same distance from the source node to the same level. Furthermore, the optimal power allocation assigns equal powers to the nodes that belong to the same level. Intuitively, this is obvious due to the symmetry in the network topology and the properties of the pathloss attenuation function. The optimal power allocation can be simplified further for such networks, thanks to the next lemma.

Lemma 2: Consider a network with channel matrix \( \mathbf{H} \). Let \( S_k \) be the set of level-\( k \) nodes. If the optimal power allocation policy assigns equal powers \( P_k \) to nodes in the same level \( S_k \), then the power allocation problem simplifies to

\[
\min_{\mathbf{p}} \mathbf{p}^T \mathbf{p} \quad \text{subject to} \quad \mathcal{L}(\mathbf{H}\hat{\mathbf{S}}\mathbf{H}^T)\hat{\mathbf{S}}\mathbf{p} \geq \mathbf{b}, \quad \mathbf{p} \geq 0, \tag{4}
\]

where \( \hat{\mathbf{p}} = [\hat{P}_1 \hat{P}_2 \ldots \hat{P}_M] \), \( \hat{P}_k = |S_k|P_k \). The \((i, k)\)th element of \( \mathbf{H} \) is \( |\mathbf{H}|_{ik} = \sum_{m \in S_k} |\mathbf{H}|_{im}/|S_k| \).

Proof: See Appendix II.
We approximate dense networks with a continuum of nodes where the relay density goes to infinity. In the continuum, after the source transmission, a certain region of the network will receive sufficient signal power. This region will be called first level and it will be denoted by $A_1$, which is a disc for broadcasting (Fig. 2). We conjecture that under the optimal broadcast scheme for the continuum network the nodes in the same level transmit with equal power. This follows due to the following reasons: (i) the pathloss attenuation function $\ell(r)$ is continuous, non-negative, decreasing and circularly symmetric; (ii) the network topology is symmetric w.r.t. source location and, hence, the nodes at the same distance from the source should behave identically. Under these conditions, each level $A_k$ becomes a thin disc. In the continuum, the transmission power will be replaced by the power density $p(r)$, which is power per unit area. Define the function

$$H(r, u) \triangleq \frac{1}{2\pi} \int_0^{2\pi} \ell(\sqrt{r^2 + u^2 - 2ur \cos(\theta)}) d\theta.$$  \hspace{1cm} (5)

Note that $H(r, u)$ represents the effective channel gain at a distance $r$ due to transmission of nodes located at a distance $u$ from the source in the continuum network.

Theorem 1: Consider the continuum network $\mathbb{S} \triangleq \{(x, y) : x^2 + y^2 \leq R^2\}$ with the source located at the center. Let $\ell(r)$ denote the pathloss attenuation function that satisfies the assumptions a1)-a3). Assume that $\ell(r)$ is such that

a1) the function $H(r, u)$ (see Eqn. 5) is decreasing in $r$ and increasing in $u$ for $0 \leq u \leq r \leq R$.

Then, the optimal power density $p(r)$ can be found as the unique continuous solution of

$$\frac{\tau \ell(r)}{\ell(0)} + \int_0^r K(r, u)p(u)du = \tau, \hspace{0.5cm} \forall r \leq R$$ \hspace{1cm} (6)

where

$$K(r, u) = H(r, u)2\pi u.$$ \hspace{1cm} (7)

Proof: See Appendix III. \hfill \blacksquare

Remark 1: Theorem 1 and Lemma 3 also apply to linear networks (see Fig. 4) where $K_1(r, u) = \ell(r - u)$ and $K_2(r, u) = \ell(r - u) + \ell(r + u)$ for these configurations.

**Example 1:** In this example, we consider the pathloss model $\ell(r) = 1/(1 + r^2)$ which is the free-space model for large $r$, $\ell(r) \approx 1/r^2$, and for small $r$, the model limits the received power at a distance $r$, $P_l \ell(r)$ to the transmit power $P_l$. Under the given pathloss model,

$$K(r, u) = \frac{2\pi u}{\sqrt{(r^2 + u^2 + 1)^2 - 4r^2u^2}},$$

$$\frac{\partial K(r, u)}{\partial r} = \frac{-4\pi ur^2 - 2 - 1}{((r^2 + u^2 + 1)^2 - 4r^2u^2)^{3/2}}.$$ \hspace{1cm} (8)

In Fig. 5, we plot the optimal power density which we evaluated by utilizing the recursive formulation (8-10). Using Lemma 4, we derive $G(r) = \pi \ln((1 + \sqrt{1 + 4\pi^2})/2)$ and $^4$Pathloss model $\ell(r) = 1/(1 + r^2)$ violates the assumption a4 when $u \leq r \leq \sqrt{a^2 + 1}$. We think that this does not affect the asymptotic behavior of $p(r)$. Furthermore, analysis can be extended to pathloss models $\ell(r) = 1/(\alpha + r^2)$, $\alpha > 0$ easily under which the asymptotic behavior is the same and the region where a4 violated is smaller for small values of $\alpha$. \hfill \blacksquare
\( \gamma = 1 \), then the asymptotic behavior of optimal power density for the pathloss model \( \ell(r) = 1/(1 + r^2) \) as \( r \to \infty \) is given by

\[
p(r) \approx \frac{\tau}{\pi \ln(r)}. \tag{12}
\]

In Fig. 5, we also plot the limiting power density (12). Furthermore, the total power transmitted by the entire network is approximately equal to

\[
P_T \approx \int_2^R \frac{\tau}{\pi \ln(r)} 2\pi r dr \approx \frac{\tau R^2}{\ln(R)}, \tag{13}
\]

for large networks, i.e., as \( R \to \infty \).

**Remark 2:** The analysis provided through continuum approximation allows us to draw conclusions for networks with finite number of nodes. Consider a finite network with node density \( \rho \). Let \( D_r \) denote the infinitesimal disc at a distance \( r \) from the source. The optimal power for a relay located at the distance \( r \) from the source can be approximated as

\[
P_{opt}(r) = \frac{\text{total relay power in } D_r}{\text{number of nodes in } D_r} = \frac{p(r) 2\pi r dr}{\rho 2\pi r dr} = \frac{p(r)}{\rho}. \tag{14}
\]

We argue that in the high density asymptote, the most power efficient scheme allows the nodes to transmit in the order of their distances from the source with power \( p(r)/\rho \).

**Remark 3:** Note that the optimal power control policy can be implemented in a distributed fashion if the nodes know their own locations. In general, OCB needs a central control unit, which requires the knowledge of all the link gains in order to schedule the transmissions. It is interesting that this is not necessary in dense networks.

**Remark 4:** The OCB allows nodes to transmit in smaller groups (most scenarios one-by-one) in order to increase the number of receptions at any node; hence, in this way the total power consumption is decreased. The main drawback of OCB is its low spectral efficiency. This is actually consistent with the fundamental result by Verdu [19] who showed that the maximum energy efficiency is achieved when the spectral efficiency is close to zero.

### VI. Conclusion

The optimal power allocation problem for cooperative broadcast is of high complexity, and it is hard to draw conclusions for general network topologies and channel models. In this paper, we approximated dense networks with a continuum of nodes which allowed us to determine the behavior of optimal power density and total power expenditure for dense large-scale networks.

**Acknowledgements**

We would like to thank Wee-Peng Tay for informing us about an error in an earlier version.

**Appendix I**

**Proof of Lemma 1**

**Proof of a):** Without loss of generality, we will assume \( \mathcal{L}(H) \) is column-ordered. We claim that for any given permutation matrix \( S \) and \( \tilde{p} \geq 0 \), \( \mathcal{L}(SHS^T)Sp \geq \tau b \Rightarrow \mathcal{L}(H)\tilde{p} \geq \tau b \). This statement means that for a given \( S \), if \( \tilde{p} \) is in the feasible set of \( p \)'s of the optimization problem

\[
\min_p 1^T p \quad \text{subject to } \mathcal{L}(SHS^T)Sp \geq \tau b, \quad p \geq 0, \tag{15}
\]

then, \( \tilde{p} \) is also in the feasible set of the optimization problem

\[
\min_p 1^T p \quad \text{subject to } \mathcal{L}(H)p \geq \tau b, \quad p \geq 0. \tag{16}
\]

This implies that the feasible set for the linear program (15) is a subset of the feasible set for the linear program (16); hence, the optimal solution is obtained by setting \( S = I \) (this is under the assumption \( \mathcal{L}(H) \) is column-ordered). The theorem follows easily. Next, we prove the claim.

Let \( E_{ij} \) denote the elementary matrix obtained by exchanging the \( i \)th and \( j \)th rows of identity matrix. It is well-known that any permutation matrix \( S \) can be written as a multiplication of elementary matrices, which is not unique. We will obtain the elementary matrices as follows. Consider the first column of the matrix \( SHS^T \). Without loss of generality assume that largest value in the first column is the first element. Determine the row operations which sorts the first column via the algorithm that shifts the \( k \)th largest value to the \((N-k+2)\)th position, starting with \( k = 2 \). Denote the elementary matrix that is associated with the \( k \)th row operation by \( E_{ik,j} \).

Then, \( S = E_{i1,j1}E_{i2,j2} \ldots E_{iN,jN} \). For simplicity, we assume that \( S = E_{ij} \) and \( 1 < i < j \leq N \). The generalization of the proof is straightforward. Let \( H_i = [H_{i,1} H_{i,2} \ldots H_{i,N+1}] \) denote the \( i \)th row of \( H \). Since, \( \mathcal{L}(H) \) is column-ordered, \( H_{i,k} \geq H_{j,k} \) for \( k < i \). Let \( \tilde{p} = [\tilde{p}_1 \ldots \tilde{p}_{N+1}] \). We know that \( \mathcal{L}(SHS^T)\tilde{p} \geq \sum_{k=1}^{i-1} H_{i,k}\tilde{p}_k \geq \tau \). We will use this condition to prove \( \mathcal{L}(H)\tilde{p} \geq \tau \), and \( \mathcal{L}(H)\tilde{p} \geq \tau \).

The \((i-1)\)th element of \( \mathcal{L}(H)\tilde{p} \) is

\[
[\mathcal{L}(H)\tilde{p}]_{i-1} = \sum_{k=1}^{i-1} H_{i,k}\tilde{p}_k \geq \sum_{k=1}^{i-1} H_{j,k}\tilde{p}_k \geq \tau. \tag{17}
\]

\(^5\)Wee-Peng Tay is associated with Laboratory for Information and Decision Systems (LIDS), Massachusetts Institute of Technology, Cambridge, MA 02139, USA.
The $(j-1)$th element of $\mathcal{L}(H)p$ is

$$\mathcal{L}(H)p_{j-1} = \sum_{k=1}^{j-1} H_{j,k}p_{k} \geq \sum_{k=1}^{i-1} H_{j,k} \tilde{p}_{k} \geq \tau. \quad (18)$$

For $k \notin \{i-1, j-1\}$, $\mathcal{L}(H)p_k = [\mathcal{L}(ShS^T)S\tilde{p}]_k$; hence, using (17) and (18), we conclude that $\mathcal{L}(H)p \geq \tau b$. 

**Proof of b):** Since $\mathcal{L}(ShS^T)$ is column-ordered, $S$ corresponds to the optimal schedule and the optimal power allocation can be found as $p_{opt} = S^T p^*$, where $p^*$ is the solution of

$$\min_{p} 1^T p \text{ subject to } \mathcal{L}(ShS^T)p \geq \tau b, \ p \geq 0. \quad (19)$$

Next, we will use proof by contradiction. We assume that $\mathcal{L}(ShS^T)$ is row-ordered and the optimal power vector $p^* = [P_1^* \ldots P_{N+1}^*]$ is such that $\mathcal{L}(ShS^T)p^* > \tau b$. For simplicity, assume that the first row satisfies with strict inequality, i.e., $[\mathcal{L}(ShS^T)p^*_1] > \tau$. Define $\beta := \max_{i,j} \frac{\alpha_{ij}}{\alpha_{ii}}$, where $\alpha_{ij}$ is the $(i,j)$th element of $\mathcal{L}(ShS^T)$. Since $\mathcal{L}(ShS^T)$ is row-ordered, $\beta < 1$ (assuming $\exists i$ such that $\alpha_{ii} \neq \alpha_{i2}$). Define $p^{**} := [P_1^* - \epsilon, P_2^* + \beta \epsilon, P_3^*, \ldots, P_{N+1}^*]$ for some $\epsilon > 0$. Since $1^T p^{**} < 1^T p^*$, $\mathcal{L}(ShS^T)p^{**} \geq \tau b$, and $p^{**} \geq 0$, $p^*$ cannot be optimal solution. Hence, the proof follows by contradiction. The proof can be easily generalized if $[\mathcal{L}(ShS^T)p^*_k] > \tau$.

**APPENDIX II**

**PROOF OF LEMMA 2**

Assume that the optimal scheduling is such that there exist at least a level that has more than one node. We assume that nodes do not transmit and receive at the same time. In this case, each node in a given level receives the transmission of previously scheduled nodes with the same power. That is, $\forall i,j \in S_m$, $\sum_{k=0}^{m-1} \sum_{n \in S_k} H_{in}P_n = \sum_{k=0}^{m-1} \sum_{n \in S_k} H_{jn}P_n$. This can be proved by contradiction. Assume that there exists $i,j \in S_m$, such that the equality is untrue; assume that node $j$ receives the previous transmissions with higher power. In this case, by assigning node $j$ to transmit before node $i$, we can decrease the total power expenditure. The proof of the lemma follows from the equality and the assumption that the nodes in a given level transmit with the same power by using simple matrix manipulations.

**APPENDIX III**

**PROOF OF THEOREM 1**

Let $P_l$ denote the optimal transmission power of the source. In the continuum, under the pathloss function $\ell(r)$ that satisfies the assumptions a1)-a3), each level is a thin disc shaped region. We conjecture that the optimal power allocation allocates equal powers to the nodes in the same level. Consider the $r$th (located at a distance $r$) and $k$th (located at a distance $u$) levels. By using Lemma 2, we obtain the simplified formulation (4). Furthermore, in the continuum,

$$H_{ik} = H(r, u) = \int_{0}^{2\pi} \frac{1}{2\pi} \ell(\sqrt{r^2 + u^2 - 2ru \cos(\theta)}) d\theta.$$


