

Minimum Energy per Bit for Wideband Wireless Multicasting: Performance of Decode-and-Forward

Aman Jain

Department of Electrical Engineering
Princeton University
Princeton, New Jersey (USA) 08540
Email: amanjain@princeton.edu

Sanjeev R. Kulkarni

Department of Electrical Engineering
Princeton University
Princeton, New Jersey (USA) 08540
Email: kulkarni@princeton.edu

Sergio Verdú

Department of Electrical Engineering
Princeton University
Princeton, New Jersey (USA) 08540
Email: verdu@princeton.edu

Abstract—We study the minimum energy per bit required for communicating a message to all the destination nodes in a wireless network. The physical layer is modeled as an additive white Gaussian noise channel affected by circularly symmetric fading. The fading coefficients are known at neither transmitters nor receivers.

We provide an information-theoretic lower bound on the energy requirement of multicasting in arbitrary wireless networks as the solution of a linear program.

We study the broadcast performance of *decode-and-forward* operating in the non-coherent wideband scenario, and compare it with the lower bounds. For arbitrary networks with k nodes, the energy requirement of decode-and-forward is within a factor of $k - 1$ of the lower bound regardless of the magnitude of channel gains. We also show that decode-and-forward achieves the minimum energy per bit in networks that can be represented as *directed acyclic graphs*, thus establishing the exact minimum energy per bit for this class of networks. We also study *regular networks* where the area is divided into cells, each cell containing at least \underline{k} and at most \bar{k} nodes placed arbitrarily within the cell. A path loss model (with path loss exponent $\alpha > 2$) dictates the channel gains between the nodes. It is shown that the ratio between the upper bound using decode-and-forward based flooding, and the lower bound is at most a constant times $\bar{k}^{\alpha+2}/\underline{k}$.

I. INTRODUCTION

Multicasting, i.e., the task of communicating the same message to multiple destinations in a communication network is an important and well studied problem. For wireline networks, the problem of energy efficient communication can be formulated as the well-known *minimum cost spanning tree* problem. However, in wireless networks, not only is there an inherent *wireless multicast advantage* [16] that allows all the nodes within the coverage range to receive the message at no additional cost (see also [2], [11]), but even distant nodes are able to hear low power, unreliable transmissions which can be used to decrease the transmission energy costs in the network. This has been termed *cooperative wireless advantage* in [6]. Therefore, the physical layer can offer energy efficiency

This research was supported in part by the U.S. Office of Naval Research under grant number N00014-07-1-0555, the U.S. Army Research Office under grant number W911NF-07-1-0185, and through collaborative participation in the Communications and Networks Consortium sponsored by the U.S. Army Research Laboratory under the Collaborative Technology Alliance Program, Cooperative Agreement DAAD19-01-2-0011. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation thereon.

benefits which need to be properly studied and exploited. This is especially important in those wireless networks, such as sensor networks [1], where the energy budget is a paramount constraint.

Prior work [6], [13], [12], [8] has shown improvement in the energy efficiency for wireless multicasting by proposing multi-hop decode-and-forward schemes which let nodes accumulate the message energy by overhearing several low power transmissions. The question of optimal performance of such schemes can be formulated as *optimal cooperative broadcast* problem [6], [13] or *accumulative broadcast* problem [12]. In these formulations, first an optimal transmission order for the nodes is chosen. Given such an order, total transmission power is minimized by solving a linear program for the power distribution over nodes, under the condition that the total power at each destination node exceeds a threshold.

Minimum energy per bit is a fundamental information-theoretic concept that quantifies the minimum cost of reliably transmitting a large message over a noisy communication channel. In our system model, we use ‘energy’ to refer to the sum total of the transmission energies at various nodes. Therefore, minimum energy per bit is the total energy cost incurred in communicating a message to all the destination nodes in the network, normalized by the size of the message. For a simple point-to-point channel with additive Gaussian noise with or without fading and with coherent or non-coherent reception, the minimum energy per bit is given by [14]

$$E_{b\min} = \frac{N_0 \log_e 2}{G} \quad (1)$$

where N_0 is the noise spectral density and G is the ‘channel gain’ or the power gain. Attaining minimum energy per bit requires infinite bandwidth (or vanishing spectral efficiency).

In this work, we study the minimum energy per bit for multicasting in wireless networks using the description of the network at the physical layer only. No restriction is placed on the nature or complexity of communication strategies that the nodes might use, and the nodes can act as relays for each other. In fact, as is shown later, clever cooperation between nodes is crucial to improve the energy efficiency of the network. We use information-theoretic techniques to give both upper and lower bounds on the minimum possible energy consumption. Our aim is to compare the broadcasting performance of decode-

and-forward operating in the non-coherent wideband scenario, with the lower bound.

In [8] (see also [9] for details), another lower bound on the minimum energy per bit for multicasting is provided in terms of the *effective network radius*, for a similar system model. This lower bound takes the same form as the right hand side of (1) with G replaced by the effective radius of the network. The bound in this paper is stronger than the effective network radius bound. In fact, the effective network radius bound can be directly derived from our bound (see Remark 3). Furthermore, most of the lower bounds given in this paper cannot be directly obtained from the effective radius bound.

Another difference from [8] is that in this work, we encompass networks of arbitrary size. In particular, we compare the performance of decode-and-forward with the lower bound for the following three different cases of wireless networks: arbitrary networks, directed acyclic graphs, and regular networks. The case of arbitrary networks illustrates the fact that decode-and-forward is not arbitrarily worse than the optimal communication scheme for any wireless network. In fact, decode-and-forward is an optimal communication scheme for a wide class of wireless networks, i.e., those networks which can be represented as directed acyclic graphs. Examples of such networks are relay networks, broadcast channels, and layered networks. Finally, by showing the (almost) order optimality of decode-and-forward in regular networks, we make the case that decode-and-forward can be feasible in many real-world networks.

In Section II, we introduce the channel model which is similar to the one used in [8]. The description of the model used for the physical layer is given in Section II-A while the notion of minimum energy per bit is defined in Section II-B. In Section III, a lower bound is given on the minimum energy per bit for multicasting in arbitrary networks. This lower bound holds for arbitrary destination sets of nodes, and is formulated as a linear program. Furthermore, the bound is shown to depend only on the channel gains between different node pairs, and not on any other property of fading distribution.

In Section IV, we focus on broadcasting rather than the general multicast setting. We analyze decode-and-forward based wideband communication schemes. Similar to the point-to-point case, these communication schemes also require unlimited bandwidth to achieve minimum energy per bit but no knowledge of the channel states at the receivers.

In Section IV-A, we show that in any network with k nodes the energy requirement of decode-and-forward is at most $k-1$ times the minimum energy per bit. Note that this factor of $k-1$ is independent of the magnitude of the channel gains between nodes. In Section IV-B, we show that decode-and-forward achieves the lower bound on the minimum energy per bit for the class of networks that can be represented as *directed acyclic graphs*. Cooperation between nodes is crucial to attain maximal energy efficiency.

Section IV-C studies *regular networks*, in which the geographical area is divided into square cells with side length s . Each cell is restricted to contain at least \underline{k} and at most \bar{k} nodes, otherwise the nodes are free to be placed anywhere within

the cells. A path loss model is imposed on the channel gains between two nodes separated by a distance of r , such that the channel gain falls off like $r^{-\alpha}$ for $\alpha > 2$ for large enough r . In order to prevent the gains from being unbounded as $r \rightarrow 0$, a restriction is imposed on the maximum gain possible. It is shown that a decode-and-forward based flooding scheme has an achievable energy per bit which is within a constant factor of $\bar{k}^{\alpha+2}/\underline{k}$ of the lower bound given in Section III. When the number of nodes within each cell is a constant, this multiplicative gap is at most a constant that does not depend on the total number of nodes or the size of a cell. The gap is poly-logarithmic (in the total number of nodes) for large random networks. The case of regular networks serves to illustrate the fact that the lower bounds and the decode-and-forward based achievable bounds are close to each other for many networks of practical interest.

II. SYSTEM MODEL

A. Channel model

We deal with a discrete-time complex additive Gaussian noise channel with fading. Suppose that the k nodes in the network are numbered $1, \dots, k$. Let node $i \in \{1, \dots, k\}$ transmit $x_{i,t} \in \mathbb{R}$ at time t , and let $y_{j,t} \in \mathbb{C}$ be the received signal at any other node $j \in \{1, \dots, i-1, i+1, \dots, k\}$. The relation between $x_{i,t}$ and $y_{j,t}$ is given by

$$y_{j,t} = \sum_{i=1}^k h_{ij,t} x_{i,t} + z_{j,t} \quad (2)$$

where, $z_{j,t}$ is circularly symmetric complex additive Gaussian noise at the receiver j , distributed according to $\mathcal{CN}(0, N_0)$. The noise terms are independent for different receivers as well as for different times. The fading between any two distinct nodes i and j is modeled by complex-valued circularly symmetric random variables $h_{ij,t}$ which are i.i.d. for different times. We also assume that $h_{ii,t} = 0$ for all nodes i and all times t . Also, for all $(i, j) \neq (\ell, m)$, the pair $h_{ij,t}$ and $h_{\ell m,t}$ is independent for all times t . *Absence of channel state information at a transmitter i* implies that $x_{i,t}$ is independent of the channel state realization $\mathbf{h}_{i,t} = (h_{i1,t}, h_{i2,t}, \dots, h_{ik,t})$ from node i to all other nodes. *Channel state information at a receiver i* implies that the channel state realization $(h_{1i,t}, h_{2i,t}, \dots, h_{ki,t})$ is available at receiver i at time t . The quantity $g_{ij} \triangleq \mathbb{E}[|h_{ij}|^2]$ is referred to as the *channel gain* between nodes i and j .

In a network with k nodes, let node 1 be the source node. Suppose that only a subset $\mathcal{R} \subseteq \{2, \dots, k\}$ (the *destination set*) of the non-source nodes is interested in receiving the message from the source node. In a *multicast* setting, \mathcal{R} contains two or more nodes. When \mathcal{R} consists of all non-source nodes, it is called a *broadcast* setting.

B. Minimum energy per bit

All the nodes in the network can act as relays and cooperate with each other in order to communicate a common message to the destination nodes. A network-wide error is declared whenever a destination node fails to decode the message correctly

by the end of the protocol. The energy consumption of the network is the sum of average transmission energy expended at all the nodes. The condition of reliable communication requires that the probability of error goes to zero for large message sizes.

Consider a code for the network with block length n . The codeword at any node i is n symbols long, denoted by $x_i^{(n)} = (x_{i,1}, x_{i,2}, \dots, x_{i,n}) \in \mathbb{C}^n$. If the message set at the source node is $\mathcal{M} = \{1, 2, \dots, M\}$, then the codeword $x_1^{(n)}(m)$ at node 1 is determined by the message m chosen equiprobably from the message set \mathcal{M} . At any other node i , the codeword $x_i^{(n)}$ is a function of the first $t-1$ inputs at the node, i.e., $x_{i,t} = x_{i,t}(y_{i,1}, \dots, y_{i,t-1})$. At each non-source node i , there may be a *decoding function* (depending on whether the node is interested in receiving the message) which decodes a message $\hat{m}_i \in \mathcal{M}$ based on the n channel outputs $(y_{i,1}, \dots, y_{i,n})$ at the node. The probability of error of the code is defined as

$$P_e \triangleq \frac{1}{M} \sum_{m \in \mathcal{M}} P[\exists i \in \mathcal{R} : \hat{m}_i \neq m | m \text{ is the message}] \quad (3)$$

Next, we define the energy per bit of the code. Let E_{total} be the expected total energy expenditure (for all nodes) of the code, i.e.,

$$E_{\text{total}} \triangleq \mathbb{E} \left[\sum_{i=1}^k \sum_{t=1}^n |x_{i,t}|^2 \right] \quad (4)$$

The *energy per bit* of the code is defined to be $E_b \triangleq E_{\text{total}} / \log_2 M$. Note that the expectation in (4) is over the set of messages at the source node, and the noise and fading realizations which affect the channel outputs (and hence, the channel inputs) at all the non-source nodes.

An $(n, M, E_{\text{total}}, \epsilon)$ code is a code over n channel uses, with M messages at the source node, expected total energy consumption at most E_{total} and a probability of error at most $0 \leq \epsilon < 1$.

Definition: (c.f. [15]) Given $0 \leq \epsilon < 1$, $E \in \mathbb{R}_+$ is an ϵ -achievable energy per bit if for every $\delta > 0$ and all sufficiently large M , an $(n, M, (E + \delta) \log_2 M, \epsilon)$ code exists. E_b is an achievable energy per bit if it is ϵ -achievable energy per bit for all $0 < \epsilon < 1$, and the *minimum energy per bit* $E_{b\min}$ is the infimum of all the achievable energy per bit values.

When energy consumption (rather than bandwidth) is a primary concern in wireless networks, $E_{b\min}$ is a sensible measure of how much energy is required to transmit a given (large enough message).

III. LOWER BOUND ON THE MINIMUM ENERGY PER BIT

In Theorem 1, we present an information-theoretic lower bound on the minimum energy per bit for arbitrary wireless networks with an arbitrary destination set. Theorem 1 holds whenever the channel states are not known at the transmitters, regardless of the channel state knowledge at the receivers. This bound is a generalization of the lower bound given in [8, Theorem 1] (see also, [9, Theorem 1]).

As before, consider a network with the node set $\{1, 2, \dots, k\}$ where node 1 is the source node and $\mathcal{R} \subset \{2, \dots, k\}$ is the

destination set of size $r = |\mathcal{R}|$. A *cut*, say u , is a subset of nodes that contains the source node such that at least one of the nodes in \mathcal{R} does not belong to u , i.e., $u \subset \{1, 2, \dots, k\}$ such that $1 \in u$ and $\mathcal{R} \not\subset u$.

Let the collection of all such cuts be $U = \{u_1, u_2, \dots, u_{|U|}\}$, where $|U| = 2^{k-r-1}(2^r - 1)$. Based on U , construct a matrix \mathbf{L} of size $|U| \times k$ as follows. The (i, j) th coefficient of \mathbf{L} is defined to be $\sum_{\ell \in u_i^c} g_{j\ell}$ if $j \in u_i$ and 0 if $j \notin u_i$. Further, let $\mathbf{1}_l$ and $\mathbf{0}_l$ be column vectors of size l containing all 1 and all 0 entries respectively.

Theorem 1. *The minimum energy per bit for multicasting is lower bounded by the value E_1 of the following linear program.*

$$\begin{aligned} E_1 \triangleq \min_{\mathbf{q} \in \mathbb{R}^k} & \mathbf{1}_k^T \mathbf{q} : \\ & \mathbf{L} \mathbf{q} \geq (N_0 \log_e 2) \mathbf{1}_{|U|} \\ & \mathbf{q} \geq \mathbf{0}_k \end{aligned} \quad (5)$$

To prove Theorem 1, we need Lemma 1 below, but first we introduce some notation. For a subset of nodes $u \subset \{1, \dots, k\}$, we use $\mathbf{y}_u \in \mathbb{C}^k$ to denote the observation vector at the nodes belonging to u , i.e., the i th element of \mathbf{y}_u is given by

$$(y_u)_i = \begin{cases} y_i & \text{if } i \in u \\ 0 & \text{if } i \notin u \end{cases}$$

Similarly, we use $\mathbf{x}_u \in \mathbb{C}^k$ to denote the transmission random vector at the set of nodes u . Without subscripts, \mathbf{y} and \mathbf{x} denote the observations and transmissions at all the nodes. Let \mathbf{H} be the random matrix formed by the fading coefficients h_{ij} , for $i, j = 1, \dots, k$. The (i, j) th coefficient of the fading random matrix \mathbf{H}_u is given by

$$(h_u)_{ij} = \begin{cases} h_{ij} & \text{if } j \in u \\ 0 & \text{otherwise} \end{cases}$$

Dropping the time indices, we can rewrite (2) to give the observation vector at the set $u^c \triangleq \{1, \dots, k\} \setminus u$ as

$$\mathbf{y}_{u^c} = \mathbf{H}_{u^c}^T \mathbf{x} + \mathbf{z}_{u^c} \quad (6)$$

where \mathbf{z}_{u^c} is the noise vector denoting noise only at the set of nodes u^c , i.e., $(z_{u^c})_i = 0$ if $i \notin u^c$.

Lemma 1. *The minimum energy per bit for multicasting is lower bounded by*

$$E_2 \triangleq \inf_{\substack{p_1, \dots, p_k \geq 0: \\ \sum_{i=1}^k p_i > 0}} \max_{u \in U} \sup_{\substack{P_{\mathbf{x}}: \\ \mathbb{E}[|x_i|^2] \leq p_i \\ \text{for } i=1, \dots, k}} I(\mathbf{x}_u; \mathbf{y}_{u^c} | \mathbf{x}_{u^c}, \mathbf{H}) \quad (7)$$

Because of space limitations, in lieu of proof, we offer a brief justification for Lemma 1 (see also [9, Lemma 1]). For the given power constraints p_1, p_2, \dots, p_k , choose any cut $u \in U$. By the *max-flow min-cut* bound [5, Theorem 15.10.1], the rate of reliable communication from the source to any destination node in the set u^c cannot exceed

$$C_u(p_1, \dots, p_k) \triangleq \sup_{\substack{P_{\mathbf{x}}: \\ \mathbb{E}[|x_i|^2] \leq p_i \\ \text{for } i=1, \dots, k}} I(\mathbf{x}_u; \mathbf{y}_{u^c} | \mathbf{x}_{u^c}, \mathbf{H}) \quad (8)$$

bits per channel use. Therefore, the number of channel uses per bit is $1/C_u$ which implies that the total energy spent per bit in communicating to one or more destination node(s) in u^c is at least $\sum_{i=1}^k p_i/C_u$. This is clearly a lower bound on the minimum energy per bit for multicasting to all nodes in \mathcal{R} , since the set of destination nodes in u^c is a subset of \mathcal{R} .

Proof of Theorem 1: It is sufficient to show that $E_1 \leq E_2$. To do so, we first need to simplify (7) according to our channel model.

Fix the power constraints $p_1, \dots, p_k \geq 0$, such that $\sum_{i=1}^k p_i > 0$, and also fix a cut $u \in U$. For a given probability distribution of \mathbf{x} (independent of \mathbf{H}), we can bound the mutual information in (7) by

$$I(\mathbf{x}_u; \mathbf{H}_{u^c}^T \mathbf{x} + \mathbf{z}_{u^c} | \mathbf{x}_{u^c}, \mathbf{H}) = I(\mathbf{x}_u; \mathbf{H}_{u^c}^T \mathbf{x}_u + \mathbf{z}_{u^c} | \mathbf{x}_{u^c}, \mathbf{H})$$

$$\leq \mathbb{E} \left[\log_2 \det \left(\mathbf{I} + \frac{1}{N_0} \text{cov}(\mathbf{H}_{u^c}^T \mathbf{x}_u | \mathbf{x}_{u^c}, \mathbf{H}) \right) \right] \quad (9)$$

$$\leq \frac{\log_2 e}{N_0} \mathbb{E} \left[\text{tr}(\text{cov}(\mathbf{H}_{u^c}^T \mathbf{x}_u | \mathbf{x}_{u^c}, \mathbf{H})) \right] \quad (10)$$

$$= \frac{\log_2 e}{N_0} \sum_{j \in u^c} \mathbb{E} \left[\left| \sum_{i \in u} h_{ij} x_i \right|^2 \right] \quad (11)$$

$$\leq \frac{\log_2 e}{N_0} \sum_{i \in u} \sum_{j \in u^c} g_{ij} p_i \quad (12)$$

where (9) is an upper bound on the mutual information under AWGN; (10) is due to Hadamard's inequality and to the fact $\log(1+x) \leq x$; and, (12) is obtained by maximizing the right hand side of (11) among all \mathbf{x} independent of $\mathbf{h}_j = (h_{j1}, h_{j2}, \dots, h_{jk})$ such that $\mathbb{E}[|x_i|^2] \leq p_i$, taking into account that the channel coefficients are independent with zero mean.

Substituting (12) in (7) we get $E_2 \geq E_3$, where

$$E_3 \triangleq \inf_{\substack{p_1, p_2, \dots, p_k \geq 0 \\ \sum_{i=1}^k p_i > 0}} \left(\sum_{i=1}^k p_i \right) \frac{N_0 \log_e 2}{\min_{u \in U} \sum_{i \in u} \sum_{j \in u^c} g_{ij} p_i} \quad (13)$$

Note that E_3 is a lower bound on the minimum energy per bit. Although it can be shown that the formulations (5) and (13) are essentially the same, i.e., $E_1 = E_3$, in order to prove Theorem 1, we only need to show $E_1 \leq E_3$.

Note from (13) that multiplying $\mathbf{p} = (p_1, \dots, p_k)$ by an arbitrary real number $\gamma > 0$ does not change the value of the right hand side of (13). Hence, the range of \mathbf{p} can be constrained to lie within a compact set. Therefore, the infimum in (13) can be replaced by a minimum. Suppose the optimum power vector in (13) that achieves E_3 is $\mathbf{p}^* = (p_1^*, \dots, p_k^*)$. Define

$$u^* \triangleq \underset{u \in U}{\text{argmin}} \sum_{i \in u} \sum_{j \in u^c} g_{ij} p_i^* \quad (14)$$

Next, set

$$\gamma \triangleq \frac{N_0 \log_e 2}{\sum_{i \in u^*} \sum_{j \in (u^*)^c} g_{ij} p_i^*} > 0 \quad (15)$$

Therefore, for all $u \in U$,

$$\sum_{i \in u} \sum_{j \in u^c} g_{ij} (\gamma p_i^*) \geq N_0 \log_e 2 \quad (16)$$

Therefore, all the constraints of (5) are satisfied by $\gamma \mathbf{p}^*$. Also, the values of the objective functions in (5) and (13) coincide at $\gamma \mathbf{p}^*$. Therefore, for every value of the objective function (13), the objective function (5) can take the same value. This implies that $E_1 \leq E_3$. Thus, E_1 is a lower bound on the minimum energy per bit. ■

Remark 1: The lower bound given in Theorem 1 has the following interpretation. Take any cut in the network which separates the source node from at least one of the destination nodes. The total transmission energy of the nodes on the source side of the cut should be such that the total received energy per bit at all the nodes on the destination side of the cut exceeds $N_0 \log_e 2$. Program (5) simply minimizes the energy sum given this constraint for each cut.

Remark 2: The lower bound given in [8, Theorem 1] was derived using the *max-flow min-cut* bound [4, Theorem 4] by considering cuts separating each destination node from all the remaining nodes. In Theorem 1, we generalize this idea by considering all possible cuts separating each destination node from the source node. In fact, as shown below in Corollary 1 and the remarks following it, the effective network radius based bound of [8, Theorem 1] is a direct consequence of Theorem 1. The result of Corollary 1 can be thought of as an improvement on [8, Theorem 1] due to the following different definition (17) of effective network radius.

Define a *cluster* S in \mathcal{R} to be a subset of the destination set of nodes, and a *cluster set* to be a collection of disjoint clusters in \mathcal{R} . Note that the nodes in a cluster set may not exhaust \mathcal{R} .

Define the *effective network radius* G for a cluster set \mathcal{S} as

$$G(\mathcal{S}) \triangleq \frac{1}{|\mathcal{S}|} \max_{i \in \{1, \dots, k\}} \sum_{S \in \mathcal{S} \setminus i} g_i(S) \quad (17)$$

where

$$g_i(S) \triangleq \sum_{\ell \in S} g_{i\ell} \quad (18)$$

and the cluster set $\mathcal{S} \setminus i$ denotes the subset of the cluster set \mathcal{S} obtained by removing the cluster containing node i , if such a cluster exists.

Corollary 1. *The minimum energy per bit required to transmit a message reliably to all the nodes in \mathcal{R} satisfies*

$$\frac{E_b}{N_0 \min} \geq \frac{N_0 \log_e 2}{G(\mathcal{S})} \quad (19)$$

for any cluster set \mathcal{S} in \mathcal{R} .

Proof sketch: Appendix A. ■

Remark 3: From Corollary 1, we can recover [8, Theorem 1] by constructing a cluster set \mathcal{S} such that each cluster consists of a single node from \mathcal{R} and the clusters exhaust all the nodes in \mathcal{R} . In this case, effective network radius is given by

$$G = \frac{1}{|\mathcal{R}|} \max_{i \in \{1, \dots, k\}} \sum_{j \in \mathcal{R} \setminus \{i\}} g_{ij} \quad (20)$$

The usefulness of Corollary 1 will be evident in Section IV-C where a lower bound based on the definition (20) of effective

network radius is not good enough to provide the bounds that we seek.

IV. DECODE-AND-FORWARD IN WIDEBAND

By decode-and-forward, we mean any scheme in which a node reliably decodes the common message before it transmits anything. In the discussion on decode-and-forward, it is assumed that neither the receivers nor the transmitters know the channel states. However, the transmitters do need to know the power at which they transmit.

In [14], it was shown that Gaussian point-to-point channels achieving minimum energy per bit require vanishing spectral efficiency. Likewise, in this section, we assume availability of arbitrarily large bandwidth. In particular, we assign each transmitter its own wide frequency band that is orthogonal to the bands of other transmitters. All receiver nodes listen to transmissions over all the bands. Wideband broadcasts are not affected by interference from other transmissions. It was shown in [14] that knowledge of the channel states does not reduce the minimum energy per bit requirements. Various wideband communication schemes can be constructed which let a receiver decode the message reliably as soon as the accumulated energy per bit at the receiver exceeds $N_0 \log_e 2$ regardless of the knowledge of channel states of the receiver. Here, we assume the use of one such scheme.

In this section, we assume the broadcast setup where all non-source nodes are destination nodes.

A. Performance of decode-and-forward in arbitrary networks

Theorem 2. *In an arbitrary network with k nodes, decode-and-forward has an achievable energy per bit E_{bDF} that satisfies*

$$E_{bDF} \leq (k-1)E_{b\min} \quad (21)$$

Proof: We prove Theorem 2 by constructing a decode-and-forward scheme that has an energy per bit which is within a factor of $k-1$ of the lower bound given by Theorem 1.

Suppose that the energy policy $\mathbf{q} = (q_1, \dots, q_k)$ achieves the minimum value of the linear program (5) for the given network, i.e.,

$$\frac{E_b}{N_0 \min} \geq \sum_{i=1}^k q_i \quad (22)$$

Using policy \mathbf{q} , we now construct a scheme that has an energy policy $\tilde{\mathbf{q}} \in \mathbb{R}_+^k$. The scheme involves communication over $k-1$ time slots, each slot consisting of transmission by one of the nodes which has already decoded the message. The transmission power levels at the nodes is determined by $\tilde{\mathbf{q}}$.

Choose any $\epsilon > 0$. The proposed decode-and-forward scheme is the following. Node $j(t)$ transmits the message with energy expenditure per bit $\tilde{q}_{j(t)} + \epsilon$ in time slot t , for $t = 1, \dots, k-1$. If node $j(t)$ is not able to decode the message by the end of time slot $t-1$, an error is declared and the algorithm terminates. The parameters $j(t)$ and $\tilde{q}_{j(t)}$ are recursively defined by Algorithm 1 in Fig. 1. Note that $\tilde{q}_{j(k)} = 0$, so no transmission is required in the slot k .

ALGORITHM 1: Determine $j(t)$ and $\tilde{q}_{j(t)}$

- 1) Set $S(1) = \phi$, $j(1) = 1$ and $\tilde{q}_1 = (k-1)q_1$.
- 2) For slots $t = 2, \dots, k$, update
 - $S(t) = S(t-1) \cup \{j(t-1)\}$
 - $j(t) = \operatorname{argmax}_{j \in \{1, \dots, k\} \setminus S(t)} \sum_{\ell \in S(t)} \tilde{q}_\ell g_{\ell j}$
 - $\tilde{q}_{j(t)} = (k-t)q_{j(t)}$

Fig. 1. The algorithm to determine the parameters $j(t)$ and $\tilde{q}_{j(t)}$

Since the choice of $\epsilon > 0$ is arbitrary, we have an achievable energy per bit $\sum_{t=1}^k \tilde{q}_t$. The conclusion of Theorem 2 is now immediate from (22) and Lemma 2. ■

Lemma 2. *For the given decode-and-forward scheme with $j(t)$ and $\tilde{\mathbf{q}}$ as determined by Algorithm 1, the following holds*

- 1) *The set $\{j(1), j(2), \dots, j(k)\}$ forms a permutation of $\{1, 2, \dots, k\}$.*
- 2) *$\sum_{t=1}^k \tilde{q}_t \leq (k-1) \sum_{t=1}^k q_t$.*
- 3) *Node $j(t)$ can decode the message by the end of time slot $(t-1)$ with vanishing probability of error, for $t = 2, \dots, k$.*

Proof: Appendix B. ■

Remark 4: When $k = 2$, i.e., for point-to-point channels, Theorem 2 gives tight results [14], i.e., $E_{bDF} = E_{b\min}$. Of course, decode-and-forward is trivial in this case.

Remark 5: The result of Theorem 2 is consistent with the result for the three-terminal network considered in [7] where the energy per bit achievable by decode-and-forward scheme was shown to be always within a factor of 2 of the lower bound.

Remark 6: The scheme proposed here is not the best decode-and-forward scheme possible. In [12], [6], [13], optimal decode-and-forward schemes are proposed (see the discussion in Section I).

B. Decode-and-Forward for Directed Acyclic Graphs (DAGs)

Take a directed graph (or a *digraph*) $\mathcal{G} = (V, E)$, where V is the set of vertices and E (edges) is a set of ordered pairs from V . We denote a directed edge from node i to node j by (i, j) . A directed graph is *acyclic* if there are no loops in it, i.e., for every node $i \in V$ there is no directed path from node i to itself.

A given wireless network can be represented by a directed acyclic graph $\mathcal{G}(V, E)$ if there is a one-to-one correspondence (say, given by the identity mapping) between nodes $1, \dots, k$ in the network and the vertex set $V = \{1, \dots, k\}$ such that

$$(i, j) \notin E \iff g_{ij} = 0 \quad (23)$$

Thus, there is no path from a node to itself that is over strictly positive channel gain links only. For a meaningful discussion, we assume that for every non-source node there is a path that takes us from the source node to it. This also implies that no edge points into the source.

Theorem 3. *For a wireless network represented by a directed acyclic graph \mathcal{G} , the minimum energy per bit $E_{b\min}$ for*

broadcasting is given by the value E_1 of the linear program (5).

Proof: The converse part is directly from Theorem 1. Let the minimum value of the linear program (5) be achieved by the energy policy $\mathbf{q} = (q_1, \dots, q_k)$. Next, we show how to achieve the minimum energy per bit using decode-and-forward with the energy policy \mathbf{q} .

Every DAG admits a *topological sort* (see, e.g., [3, Chapter 22.4]), which is simply an ordering of nodes from left to right such that every edge is directed from left to right. For graph \mathcal{G} , since node 1 has edges only directed away from it, it is possible to start the sort with node 1. For the sake of simplicity, we assume that all the nodes in \mathcal{G} are already topologically sorted (i.e., for every directed edge $(i, j) \in E$, $i < j$).

Choose an $\epsilon > 0$. Consider the following simple decode-and-forward scheme according to the power policy \mathbf{q} . Node t transmits the message with energy expenditure per bit $q_t + \epsilon$ in time slot t , for $t = 1, \dots, k-1$. If node t is not able to decode the message by the end of time slot $t-1$, an error is declared and the algorithm terminates.

Lemma 3 shows that the probability of error is arbitrarily small. Note that since the choice of $\epsilon > 0$ is arbitrary, we can have an achievable energy per bit of $\sum_{i=1}^k q_i$. This immediately implies that the value of the program (5) is achievable. ■

Lemma 3. *For the described decode-and-forward scheme, and for every $1 \leq t \leq k-1$, node $t+1$ successfully decodes the message by the end of time slot t with vanishing probability of error.*

Proof: Appendix C. ■

C. Regular Networks

In a regular network, $k \geq 2$ nodes are placed within a square network area under the following restrictions. If we divide the network area into n^2 square cells of side length s , then each such cell contains at most \bar{k} non-source nodes and at least $\underline{k} \geq 1$ non-source nodes. The nodes can be placed arbitrarily within cells. Let the network area be a square with the diagonal coordinates $(0, 0)$ and $(n s, n s)$. The source node, node 1, is placed at the origin. Let \mathcal{C} denote the set of cells in the network. We use $C(x, y) \in \mathcal{C}$ for $x, y = 0, 1, \dots, n-1$, to denote the cell with its lower left corner at the coordinates $(x s, y s)$. For simplicity, we restrict our attention to the case when n is a multiple of 3. Note that

$$\frac{k}{\bar{k}} \leq n^2 \leq \frac{k}{\underline{k}} \quad (24)$$

The channel gain between any pair of nodes, say i and j , is determined solely by the distance r_{ij} between the nodes through a monotonically decreasing *path loss function* $g(r) : \mathbb{R}_+ \rightarrow \mathbb{R}_+$, i.e., $g_{ij} = g(r_{ij})$, which satisfies for all $r \geq r_0$

$$g(r) = r^{-\alpha} \quad (25)$$

where $r_0 > 0$ and $\alpha > 2$ are constants of the model. We impose an additional restriction on the path loss function that

FLOOD(E_{b_1}, E_{b_2}) [8]

Let T be the total number of time slots, depending on the network parameters.

- 1) The source node transmits only in the 1st time slot with energy per bit E_{b_1} .
- 2) At the beginning of time slot $t = 2, \dots, T$, each node (except the source node) executes the following
 - If the node was able to decode a message successfully for the first time in the previous time slot, then it retransmits the same message in the current time slot with energy per bit E_{b_2} .
 - Else, keep quiet.

Fig. 2. FLOOD(E_{b_1}, E_{b_2})

$g(0) \leq \bar{g}$ for some constant $\bar{g} > 0$ to ensure that the gain between any two nodes is not arbitrarily large.

For the converse results on minimum energy per bit in regular networks, we use Corollary 1. On the other hand, achievability is demonstrated using a decode-and-forward scheme which is also a flooding algorithm at the network layer (Algorithm FLOOD given in Fig. 2). We operate under the *minimal information framework* of [8] which restricts any location or channel state information at the receivers. Also, the nodes do not follow any transmission schedule and transmit at a pre-decided energy per bit within their own wideband once they decode a message.

Theorem 4. *In a regular network, the minimum energy per bit $E_{b_{\min}}$ and the achievable energy per bit $E_{b_{\text{flood}}}$ of FLOOD satisfy*

$$\frac{E_{b_{\text{flood}}}}{E_{b_{\min}}} \leq c \frac{\bar{k}^{\alpha+2}}{\underline{k}} \quad (26)$$

where, $c \in \mathbb{R}_+$ is some constant depending only on the path loss model.

Before getting into the proof, we do some elementary constructions. Divide each cell into $(9\bar{k} + 1) \times (9\bar{k} + 1)$ square *subcells*. Also, partition the whole network into *supercells* consisting of 9 cells each. Each supercell consists of a *leader cell* of the form $C(3m_1 + 1, 3m_2 + 1)$ for $m_1, m_2 = 0, 1, \dots, (n/3) - 1$, and the 8 cells (*subordinate cells*) that are geographically tangent with the leader cell. Note that there are $n^2/9$ supercells in the network. (See Fig. 3).

Next, with each leader cell, say $C(3m_1 + 1, 3m_2 + 1)$, we associate a cluster S_{m_1, m_2} of nodes constructed by Algorithm 2 given in Fig. 4.

Applying Algorithm 2 to each leader cell gives us a cluster set $\mathcal{S} = \{S_{m_1, m_2} : m_1, m_2 \in \{0, \dots, (n/3) - 1\}\}$. A few observations about the clusters and the cluster set are given in Lemma 4 without proof.

Lemma 4. *For the cluster set \mathcal{S} generated by Algorithm 2, the following hold*

- 1) *For any cluster, the minimum distance between a node in the cluster and any node not in the cluster is $s/(9\bar{k} + 1)$.*
- 2) *The cluster S_{m_1, m_2} is a subset of nodes belonging to the supercell with the leader cell $C(3m_1 + 1, 3m_2 + 1)$.*
- 3) $|\mathcal{S}| = n^2/9$.

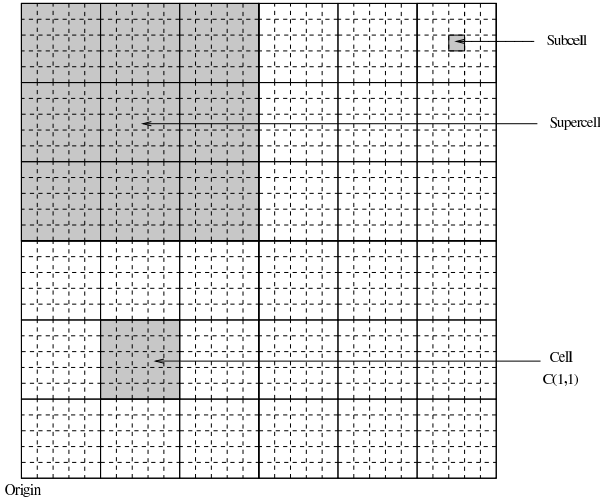


Fig. 3. A regular network with 36 cells, 25 subcells within each cell, and 4 supercells each consisting of 9 cells.

ALGORITHM 2: Construction of Cluster S_{m_1, m_2}

- 1) Initialize $S_{m_1, m_2} = \{\text{All nodes in the leader cell } C(3m_1 + 1, 3m_2 + 1)\}$.
- 2)
 - If there are nodes in any subcell adjacent (vertically, horizontally or diagonally) to the subcell to which a node in S_{m_1, m_2} belongs, then include those nodes in S_{m_1, m_2} . Repeat step 2.
 - Else, return S_{m_1, m_2} and terminate.

Fig. 4. The algorithm to construct clusters S_{m_1, m_2}

4) Each cluster has at least \underline{k} nodes and at most $9\bar{k}$.

Proof sketch for Theorem 4: Knowing S from Algorithm 2, the converse is obtained by applying Corollary 1. In order to do so, we need to upper bound the following quantity for all $i = 1, \dots, k$

$$\sum_{S \in \mathcal{S} \setminus i} g_i(S) \quad (27)$$

which is the sum of gains from node i to all the nodes which are part of some cluster not containing node i . Since each cluster has a one-to-one correspondence to leader cells and supercells, instead of summing over the clusters the quantity in (27) can be summed over the corresponding supercells. Node i belongs to a supercell, say C . Consider the supercells at ℓ steps (horizontal, vertical or diagonal) away from C , for $\ell = 0, 1, \dots, (n/3) - 1$. The supercell at a step distance of 0 is C itself. So, node i is either a part of the cluster of C or it is at a distance of at least $s/(9\bar{k} + 1)$ from all nodes in the cluster. Similarly, node i is at least $s/(9\bar{k} + 1)$ distance away from any node in the clusters corresponding to the supercells surrounding C . For $\ell \geq 2$, there are at least $(\ell - 1)$ whole supercells separating node i from a cluster in the supercell ℓ steps away, thus giving the minimum separation between those clusters and node i to be $3(\ell - 1)s$. Furthermore, from Lemma 4, there are at most $9\bar{k}$ nodes in any cluster. Also, there are at most 8ℓ clusters at a step distance of ℓ . Thus, using (17) we

get the bound on $G(S)$ as

$$G(S) \leq \frac{729 \bar{k}^2}{k} \left[g \left(\frac{s}{(9\bar{k} + 1)} \right) + \sum_{\ell=2}^{\sqrt{n}/3-1} \ell g(3(\ell - 1)s) \right] \quad (28)$$

The proof is divided into three cases:

- *Case 1:* $ns/(9\bar{k} + 1) < r_0$

The condition of this case implies that the whole network can be enclosed within a box of side $(9\bar{k} + 1)r_0$. Thus, the following single shot transmission scheme

$$\text{FLOOD} \left(\frac{N_0 \log_e 2}{g(\sqrt{2}(9\bar{k} + 1)r_0)} + \epsilon_1, 0 \right) \quad (29)$$

for any $\epsilon_1 > 0$, works well. Note that, in this case, $T = 1$ in the algorithm FLOOD.

On the other hand, since the gain to a node cannot exceed \bar{g} , the effective network radius (in the sense of [8, Theorem 1]) does not exceed \bar{g} . Therefore, by [8, Theorem 1], $E_{b_{\min}} \geq N_0 \log_e 2 / \bar{g}$. Using (25) and the fact that the energy consumption $E_{b_{\text{flood}}}$ of $\text{FLOOD}(E_{b_1}, E_{b_2})$ is at most $E_{b_1} + (k - 1)E_{b_2}$, we can evaluate the $E_{b_{\text{flood}}}$ for (29) which gives

$$\frac{E_{b_{\text{flood}}}}{E_{b_{\min}}} \leq c_1 \bar{k}^\alpha \quad (30)$$

for some positive constant c_1 .

- *Case 2:* $r_0 < s/(9\bar{k} + 1)$

Under this condition, all the distances of interest in (28) are greater than r_0 . Using (25), we can bound $G(S)$ as

$$G(S) \leq \frac{729 \bar{k}^2}{k} \left[\frac{(9\bar{k} + 1)^\alpha}{s^\alpha} + \sum_{\ell=2}^{\infty} \frac{\ell}{3^\alpha (\ell - 1)^\alpha s^\alpha} \right] \leq c_2 \frac{\bar{k}^{\alpha+2}}{k s^\alpha} \quad (31)$$

for some positive constant c_2 , where, in deriving (31) we have made use of the facts that $\ell/(\ell - 1)^\alpha \leq 2^\alpha \ell^{1-\alpha}$ for $\ell \geq 2$, and that $\sum_{\ell=1}^{\infty} \ell^{1-\alpha} < \infty$ since $\alpha > 2$. So, from Corollary 1, $E_{b_{\min}} \geq (N_0 \log_e 2) k s^\alpha / (c_2 \bar{k}^{\alpha+2})$.

Next,

$$\text{FLOOD} \left(\frac{N_0 \log_e 2}{g(2\sqrt{2}s)} + \epsilon_1, \frac{N_0 \log_e 2}{k g(2\sqrt{2}s)} + \epsilon_1 \right) \quad (32)$$

for any $\epsilon_1 > 0$, can reach all the nodes. To see this, note that the first transmission, by the source, is enough to reach all the nodes within a distance of $2\sqrt{2}s$. Thereafter, if all the nodes (at least \underline{k} in number) in a cell have already decoded a message, then their transmissions according to (32) ensure that all the nodes in an adjacent (horizontally, vertically or diagonally) cell receive enough energy ($> N_0 \log_e 2$) to decode the message with arbitrarily small probability of error. Note that, for this case, $T = n$ in the algorithm FLOOD.

Since $2\sqrt{2}s > r_0$ for this case, $E_{b_{\text{flood}}} \leq 8^{\alpha/2} s^\alpha (1 + (k -$

1)/ \bar{k}) $N_0 \log_e 2$. Therefore,

$$\frac{E_{b\text{flood}}}{E_{b\text{min}}} \leq c_3 \frac{\bar{k}^{\alpha+2}}{\underline{k}} \quad (33)$$

for some positive constant c_3 .

- *Case 3:* $s/(9\bar{k} + 1) \leq r_0 \leq ns/(9\bar{k} + 1)$

Define

$$L \triangleq \lfloor \frac{(9\bar{k} + 1)r_0}{s} \rfloor \geq 1$$

Next, G can be bounded as

$$\begin{aligned} G(\mathcal{S}) &\leq \frac{729\bar{k}^2}{k} \left[\frac{L(L+1)}{2} \bar{g} + \sum_{\ell=L+1}^{\infty} \frac{\ell}{3^\alpha(\ell-1)^\alpha s^\alpha} \right] \\ &\leq c_4 \frac{\bar{k}^4}{k s^2} \end{aligned} \quad (34)$$

for some positive constant c_4 . The details of derivation of (34) are omitted in the interest of space, but are along the lines of (31) taking care of the fact that for $\ell \leq L$, the supercells ℓ steps away may lie in the near-field of fading. So, from Corollary 1, (34) gives us $E_{b\text{min}} \geq (N_0 \log_e 2) k s^2 / (c_4 \bar{k}^4)$. Furthermore,

$$\text{FLOOD} \left(\frac{N_0 \log_e 2}{g(2\sqrt{2}sL)} + \epsilon_1, \frac{N_0 \log_e 2}{\underline{k} \sum_{\ell=1}^L \ell g(2\sqrt{2}s\ell)} + \epsilon_1 \right) \quad (35)$$

for any $\epsilon_1 > 0$, reaches all the nodes. To see this, note that after the first transmission by the source, the message is successfully received by all the nodes lying within the set of cells \mathcal{S}_1 , where

$$\mathcal{S}_t \triangleq \{C(x, y) : \max\{x, y\} \leq (L-1) + t\} \quad (36)$$

for $1 \leq t \leq n - L$. Assuming that all the nodes in the cell \mathcal{S}_t decode the message by the end of time slot t , it can be verified that all the nodes in the cells $\mathcal{S}_{t+1} \setminus \mathcal{S}_t$ would have received enough energy from the nodes (transmitting according to (35)) in the cells \mathcal{S}_t in order to decode the message with arbitrarily small probability of error. We have used the fact that for every $1 \leq \ell \leq L$, each cell in $\mathcal{S}_{t+1} \setminus \mathcal{S}_t$ has ℓ or more distinct cells in \mathcal{S}_t which are exactly ℓ vertical, horizontal or diagonal steps away. By induction, algorithm (35) is good enough to reach all the nodes in the network. From (36), it is clear that $T = n - L + 1$ for this case. Finally, the energy consumption of (35) can be shown to be $E_{b\text{flood}} \leq (N_0 \log_e 2) c_5 k s^2 \bar{k}^{\alpha-2} / \underline{k}$ for some positive constant c_5 . Therefore,

$$\frac{E_{b\text{flood}}}{E_{b\text{min}}} \leq c_4 c_5 \frac{\bar{k}^{\alpha+2}}{\underline{k}} \quad (37)$$

The final statement is now immediate by combining all the cases. ■

Remark 7: According to Theorem 4, if we bound the maximum number of nodes in each cell \bar{k} by a constant, the ratio of $E_{b\text{flood}}$ and $E_{b\text{min}}$ is at most a constant which does not depend on the number or size of cells in the network.

Remark 8: Suppose that there are $k - 1$ non-source

nodes in a network which are placed randomly uniformly and independently over a square area of size A_k . If the network area is divided into square cells of size $s_k = \sqrt{3A_k \log_e(k-1)}/\sqrt{k-1}$, then $\underline{k} \geq 1$ and $\bar{k} \leq 3e \log_e(k-1)$ almost surely, as $k \rightarrow \infty$ [10, Claim 3.1]. Therefore, for random placement of nodes, by Theorem 4, we get

$$\frac{E_{b\text{flood}}}{E_{b\text{min}}} = O((\log k)^{\alpha+2}) \quad (38)$$

V. CONCLUSION

We have studied the maximum possible energy efficiency in wireless networks using only the physical layer description. An information-theoretic lower bound on the minimum energy per bit for multicasting in arbitrary networks is derived.

We have also proposed a wideband communication scheme based on decode-and-forward. We have shown that this communication scheme is worse than the lower bound by a factor of at most $k - 1$ for broadcasting in arbitrary networks, where k is the total number of nodes in the network. This result implies that the performance of decode-and-forward cannot be arbitrarily worse than that of the optimal scheme. Furthermore, the gap is not too large for smaller networks, especially since it is unaffected by the magnitude of the channel gains.

On the other hand, decode-and-forward is shown to attain the exact minimum energy per bit in directed acyclic graphs. The optimal energy efficiency is attained by cooperation between the nodes, where the nodes collect energy over multiple time slots and the nodes that have decoded the message then retransmit it for the benefit of their peers.

Though the gap between the upper and lower bounds can be large in arbitrary networks, if the number of nodes within each cell is fixed, then the gap is independent of the network size. Furthermore, regular networks can be used to study large random networks to show that the gap is at most a poly-logarithmic factor in the number of nodes. Thus, regular networks demonstrate the utility of the given lower bound and simple decode-and-forward schemes in many practical setups of interest.

APPENDIX A PROOF SKETCH OF COROLLARY 1

Proof sketch: The proof proceeds by modifying the linear program (5) so that the number of constraints is reduced. In particular, we choose to retain those rows of \mathbf{L} (to form a new matrix $\tilde{\mathbf{L}}$) that correspond to the cut S_i^c for all $S_i \in \mathcal{S}$. Since the constraints are reduced, the value of the linear program (5) with \mathbf{L} replaced by $\tilde{\mathbf{L}}$ is less than the value of linear program (5) with \mathbf{L} . Then, consider the dual of the resulting program and evaluate its value at the specific admissible vector $\gamma^{-1} \mathbf{1}_{|\mathcal{S}|}$ where

$$\gamma = \max_{i=1, \dots, k} \sum_{S \in \mathcal{S} \setminus i} \sum_{\ell \in S} g_{i\ell} \quad (39)$$

This provides the requisite lower bound on the minimum energy per bit. ■

APPENDIX B
PROOF OF LEMMA 2

Proof: The first part of Lemma 2 is true since $j(1) = 1$ and each of the remaining $k - 1$ updates in Algorithm 1 sets a value of j which has not been considered before.

The second part of Lemma 2 follows directly from Algorithm 1 since,

$$\tilde{q}_{j(t)} \leq (k - 1) q_{j(t)} \quad (40)$$

for $t = 1, \dots, k$.

Now for the third part of Lemma 2. First, note that according to Algorithm 1, $S(t)$ is nothing but the set of nodes which have transmitted by the end of time slot $t-1$, i.e., $S(t) = \bigcup_{\ell=1}^{t-1} j(\ell)$. For $t = 1$, $j(t) = 1$ which is the source node and already has the message. Note that $S(2) = \{1\}$. For any $t \geq 2$, suppose that the nodes in $S(t)$ have all decoded the message. We claim that node $j(t)$ has gathered enough energy ($\geq N_0 \log_e 2$) by the end of slot $t - 1$ for it to decode the message reliably. Suppose not, i.e., for all $j \in S(t)^c \triangleq \{1, \dots, k\} \setminus S(t)$ we have

$$\sum_{\ell \in S(t)} \tilde{q}_\ell g_{\ell j} < N_0 \log_e 2 \quad (41)$$

This implies,

$$\sum_{j \in S(t)^c} \sum_{\ell \in S(t)} \tilde{q}_\ell g_{\ell j} < (k - t + 1) N_0 \log_e 2 \quad (42)$$

However, by our update rule for $\tilde{q}_{j(t)}$, for all $\ell \in S(t)$, $\tilde{q}_\ell \geq (k - t + 1) q_\ell$. Therefore,

$$\sum_{j \in S(t)^c} \sum_{\ell \in S(t)} q_\ell g_{\ell j} < N_0 \log_e 2 \quad (43)$$

Recall that \mathbf{q} satisfies the constraints laid out in (5). Consider the row in G corresponding to the cut $u = S(t)$. The constraint due to u dictates that

$$\sum_{\ell \in S(t)} q_\ell \sum_{j \in S(t)^c} g_{\ell j} \geq N_0 \log_e 2 \quad (44)$$

which directly contradicts (43). Therefore, there does exist a node in $S(t)^c$ for which the total received energy per bit from the nodes in $S(t)$ is sufficient to decode the message reliably. By our choice of $j(t)$ in Algorithm 1, we are sure to pick one such node. Finally, by induction, the third part of Lemma 2 is true for all $t = 2, \dots, k$. ■

APPENDIX C
PROOF OF LEMMA 3

Proof: By induction. In the first time slot, only node 1 transmits the message with energy per bit q_1 . Consider the cut isolating node 2, i.e., the cut $\{1, 3, 4, \dots, k\}$. Since this cut is valid, its corresponding constraint in (5) dictates that

$$g_{12} q_1 \geq N_0 \log_e 2 \quad (45)$$

since g_{i2} is zero for $i = 2, 3, \dots, k$. Thus, the received energy per bit at node 2 is enough for it to reliably decode the message by the end of time slot 1.

Now, for $2 \leq t \leq k - 1$, suppose that nodes $2, \dots, t$ decode the message successfully by the end of time slot $t-1$. Consider

the cut which isolates node $t + 1$, i.e., the cut $\{1, 2, \dots, t, t + 2, \dots, k\}$. This is a valid cut which appears as some row in G in (5). Therefore,

$$\sum_{\substack{i=1 \\ i \neq t+1}}^k q_i g_{i(t+1)} \geq N_0 \log_e 2 \quad (46)$$

Since we have the topological sort on the set of nodes, $g_{i(t+1)} = 0$ for $i > t + 1$. Thus, (46) is equivalent to

$$\sum_{i=1}^t q_i g_{i(t+1)} \geq N_0 \log_e 2 \quad (47)$$

Since the nodes $1, 2, \dots, t$ have already decoded the message and transmitted by the end of time slot t according to the energy policy \mathbf{q} (plus an $\epsilon > 0$), node $t + 1$ receives enough energy per bit ($> N_0 \log_e 2$) to decode the message with vanishing probability of error by the end of slot t . ■

REFERENCES

- [1] I. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *IEEE Commun. Mag.*, vol. 40, no. 8, pp. 102–114, Aug. 2002.
- [2] M. Çagalj, J. Hubaux, and C. Enz, "Minimum-energy broadcast in all-wireless networks: NP-completeness and distribution issues," in *Proc. 8th Annual International Conference on Mobile Computing and Networking (MOBICOM 2002)*, Atlanta, USA, Sep. 2002.
- [3] T. Cormen, C. Leiserson, R. Rivest, and C. Stein, *Introduction to Algorithms*, 2nd ed. Cambridge, MA, USA: MIT Press, 2001.
- [4] T. Cover and A. E. Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. 25, no. 5, pp. 572–584, Sep. 1979.
- [5] T. Cover and J. Thomas, *Elements of Information Theory*, 2nd ed. New York, NY, USA: John Wiley & Sons, 2006.
- [6] Y. Hong and A. Scaglione, "Energy-efficient broadcasting with cooperative transmissions in wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 5, no. 10, pp. 2844–2855, Oct. 2006.
- [7] A. Jain, S. R. Kulkarni, and S. Verdú, "Minimum energy per bit for gaussian broadcast channels with cooperating receivers and common message," in *Proc. Allerton Annual Conference on Communication, Control and Computing*, Monticello, IL, USA, Sep. 2009.
- [8] —, "Multicasting in large random wireless networks: Bounds on the minimum energy per bit," in *Proc. 2009 IEEE Symposium on Information Theory (ISIT 2009)*, Seoul, South Korea, Jun. 2009.
- [9] —, "Multicasting in large wireless networks: Bounds on the minimum energy per bit," May 2009. [Online]. Available: <http://jpp.arxiv.org/abs/0905.3858>
- [10] S. R. Kulkarni and P. Viswanath, "A deterministic approach to throughput scaling in wireless networks," *IEEE Trans. Inf. Theory*, vol. 50, no. 6, pp. 1041–1049, Jun. 2004.
- [11] W. Liang, "Constructing minimum-energy broadcast trees in wireless ad hoc networks," in *Proc. 3rd ACM international symposium on Mobile ad hoc networking & computing (MOBIHOC 2002)*, Lausanne, Switzerland, Jun. 2002.
- [12] I. Maric and R. D. Yates, "Cooperative multihop broadcast for wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 22, no. 6, pp. 1080–1088, Aug. 2004.
- [13] B. S. Mergen and A. Scaglione, "On the power efficiency of cooperative broadcast in dense wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 2, pp. 497–507, Feb. 2007.
- [14] S. Verdú, "Spectral efficiency in the wideband regime," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1319–1343, Jun. 2002.
- [15] —, "On channel capacity per unit cost," *IEEE Trans. Inf. Theory*, vol. 36, no. 5, pp. 1019–1030, Sep. 1990.
- [16] J. E. Wieselthier, G. D. Nguyen, and A. Ephremides, "On the construction of energy-efficient broadcast and multicast trees in wireless networks," in *Proc. 19th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2000)*, vol. 2, Tel-Aviv, Israel, Mar. 2000.