Energy Efficiency of Decode-and-Forward for Wideband Wireless Multicasting

Aman Jain, Sanjeev R. Kulkarni, *Fellow, IEEE*, and Sergio Verdú, *Fellow, IEEE*

***Abstract*—In this paper, 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 (AWGN) 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 general multicasting in arbitrary networks as the solution of a linear program, when no restrictions are placed on the bandwidth or the delay. We study the performance of *decode-and-forward* operating in the noncoherent wideband scenario, and compare it with the lower bound, for a variety of network classes where all nonsource nodes are destinations. For three-terminal networks with one source and two cooperative destination nodes, the energy expenditure of decode-and-forward is shown to be at most twice the lower bound and optimal in many cases. We also show that for arbitrary networks with k nodes, the energy requirement of decode-and-forward is at most $k-1$ times that of the lower bound regardless of the magnitude of channel gains. In networks that can be represented as directed acyclic graphs (DAGs), we establish the minimum energy per bit, also achieved by decode-and-forward. In addition, we also study *regular networks* where the energy consumption of decode-and-forward is shown to be almost order optimal in many situations of interest.**

***Index Terms*—Decode-and-forward, flooding, minimum energy per bit, multicasting, noncoherent communication, relay networks.**

I. Introduction

MULTICASTING, i.e., the problem of communicating the same message to multiple destinations in a communication network has received abundant attention in the literature. 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 [25] that allows all the nodes within the coverage range to receive the message at no additional cost (see also [3], [17]), 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 thanks to cooperation and relaying. This has been termed cooperative wireless advantage in [10]. Therefore, the physical layer can offer energy efficiencies 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, while spectral efficiency is secondary.

Prior work [10], [20], [19], [11] has shown improvement in the energy efficiency of wireless multicasting by proposing multihop 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 either as optimal cooperative broadcast [10], [20] or as *accumulative broadcast* [19]. In these formulations, 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.

In this work, we quantify the energy efficiency as minimum energy per bit $E_{b\min}$. Minimum energy per bit is a fundamental information-theoretic concept that quantifies the minimum cost of transmission over a noisy communication channel. When energy consumption (rather than bandwidth) is a primary concern, $E_{b\min}$ is a sensible measure of how much transmission energy is required for reliably transmitting asymptotically long message blocks, when there is no restriction on the number of channel uses. For a simple point-to-point channel with additive white Gaussian noise (AWGN) and fading, the minimum energy per bit is given by

$$E_{b\min} = \frac{N_0 \log_e 2}{G} \tag{1}$$

where N_0 is the noise spectral density and G is the "channel gain" (or equivalently, $1/G$ is the "channel loss") [23]. $E_{b\min}$ is not affected by the knowledge of channel states at the receiver. Attaining minimum energy per bit requires infinite bandwidth (or vanishing spectral efficiency).

Minimum energy per bit is also known for the Gaussian *multiple-access channel*, the Gaussian *broadcast channel*, and the Gaussian *interference channel* [24], [23], [15], [4]. However, it is not yet known for the three-terminal setting of a Gaussian *relay channel*, though some progress has been made (see [9], [26], and the references therein).

In this paper, we study the minimum energy per bit for multicasting in wireless networks using the description of the net-

Manuscript received November 29, 2009; revised February 07, 2011; accepted April 27, 2011. Date of current version December 07, 2011. This work was supported in part by the U.S. Office of Naval Research under Grant N00014-07-1-0555, by the U.S. Army Research Office under Grant W911NF-07-1-0185, and by the Center for Science of Information (CSoI), an NSF Science and Technology Center, under Grant CCF-0939370. The material in this paper was presented in part at the 47th Annual Allerton Conference on Communication, Control, and Computing, Monticello, IL, September 2009.

A. Jain was with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544 USA. He is now with Goldman Sachs & Co., New York, NY 10282 USA (e-mail: amanjain@princeton.edu).

S. R. Kulkarni and S. Verdú are with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544 USA (e-mail: kulkarni@princeton.edu; verdu@princeton.edu).

Communicated by R. D. Yates, Associate Editor for Communication Networks.

Digital Object Identifier 10.1109/TIT.2011.2170120

0018-9448/$26.00 © 2011 IEEE

work at the physical layer only. The physical layer is modeled as a Gaussian channel with circularly symmetric fading. The fading coefficients are known at neither transmitters nor receivers. Each transmission incurs a cost quadratic in the magnitude of the transmitted symbol. In our system model, we use "energy" to refer to the sum total of the transmission costs at all the nodes.

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 noncoherent wideband scenario with the lower bound for the following cases of wireless networks: three-terminal networks, arbitrary networks, directed acyclic graphs (DAGs), and regular networks.

In [11], another lower bound on the minimum energy per bit for multicasting is provided in terms of the effective network loss, for a similar system model. This lower bound takes the same form as the right-hand side of (1) with $1/G$ replaced by the effective loss of the network. The bounds in this paper are stronger than the effective loss bound of [11]. In fact, none of the lower bounds given in this paper can be directly obtained from the effective loss bound.

In Section II, we introduce the channel model, which is similar to the one used in [11]. 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 the fading distribution.

In Section IV, we focus on broadcasting to all the nodes rather than the general multicast setting where not all nodes serve as destinations. We analyze decode-and-forward-based wideband communication schemes. Due to the assumption of large bandwidth, each node is allocated a separate and dedicated wideband frequency channel. 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-B, the three-terminal problem with one source and two cooperating destination nodes is studied. It is shown that the energy expenditure of decode-and-forward is always within a factor of 2 of the lower bound, and in many cases, decode-and-forward achieves the minimum energy per bit. This problem has been considered before as fully cooperative relay broadcast channel (or cooperative broadcast channel, in short) with a common message (see [18], [8], [21], and references therein) for general channels. The case of Gaussian channels without fading has been studied in [18], [16], and [2].

In Section IV-C, we show that in any k-node network the energy requirement of decode-and-forward for broadcasting 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-D, 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 DAGs. Cooperation between nodes is crucial to attain this maximal energy efficiency. Popular network settings that can be represented as DAGs include relay networks, broadcast channels, and layered networks.

In Section IV-E, we study *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 unbounded gains as $r \to 0$, a restriction is imposed on the maximum possible gain. It is shown that a decode-and-forward-based flooding scheme has an achievable energy per bit within a constant factor of $\bar{k}^{\alpha+2}/\underline{k}$ of the lower bound given in Section III. This multiplicative gap is a constant when the number of nodes within each cell is a constant, and is poly-logarithmic (in the number of nodes) for large random networks (cf. [11]). The case of regular networks serves to illustrate the fact that the lower bound and the decode-and-forward upper bound are close for many networks of practical interest.

II. System Model

In this section, we present the channel model and the necessary information-theoretic background. Unless otherwise specified, vectors are represented with boldface, small-case letters (e.g., $\mathbf{h}$) and matrices are represented with boldface, upper case letters (e.g., $\mathbf{H}$) throughout the paper.

A. Channel Model

We deal with a discrete-time complex AWGN channel with fading. Let node $i \in \{1,\ldots,k\}$ transmit $x_{i,t} \in \mathbb{C}$ at time t, and let $y_{j,t} \in \mathbb{C}$ be the received signal at node $j \in \{1,\ldots,k\}$ at time t. The relation between $x_{i,t}$ and $y_{j,t}$ is given by

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

where $z_{j,t}$ is circularly symmetric complex AWGN at 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 independent identically distributed (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 $\mathbf{h}_{i,t} = (h_{i1,t}, h_{i2,t}, \ldots, h_{ik,t})$ from node i to all other nodes. Channel state information at a receiver i implies that the channel state realization $(h_{1i,t}, h_{2i,t} \ldots, h_{ki,t})$ is known at receiver i at time t. The quantity $g_{ij} \triangleq \mathbb{E}[|h_{ij}|^2] \in \mathbb{R}_+$ is referred to as the *channel gain* between nodes i and j.

For concreteness, let node 1 be the source node. Suppose that only a subset $\mathcal{R} \subseteq \{2,\ldots,k\}$ (called the *destination set*) of the nonsource nodes is interested in receiving the message from the source node. In a *multicast* setting, $\mathcal{R}$ contains two or more nodes [22, Sec. 5.2.7 and 5.2.8]. When $\mathcal{R}$ consists of all nonsource nodes, we call it 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. An 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 the average transmission energy expended at all the nodes. Reliable transmission 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 $x_i^{(n)} = (x_{i,1}, x_{i,2}, \ldots, x_{i,n}) \in \mathbb{C}^n$. If the message set at the source node is $\mathcal{M} = \{1, 2, \ldots, M\}$, 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., $x_{i,t} = x_{i,t}(y_{i,1}, \ldots, y_{i,t-1})$. At each nonsource 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 $y_i^{(n)} = (y_{i,1}, \ldots, y_{i,n})$ at the node. Therefore, $\hat{m}_i = \hat{m}_i(y_i^{(n)})$.

The probability of error of the code is defined as

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

Note that the error event at a single node is a subset of the error event defined above. Clearly, P_e is at least as large as the probability of error at any subset of the nodes in $\mathcal{R}$.

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.,

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

where $E_{i,t} \triangleq \mathbb{E}[|x_{i,t}|^2]$ is the expected energy spent in transmitting the tth symbol at node i. Note that, in each case, the expectation is over the set of messages at the source node, and the noise and fading realizations which affect the channel outputs (and hence, the relay inputs).

The energy per bit of the code is defined to be

$$E_b \triangleq \frac{E_{\text{total}}}{\log_2 M}. \quad (5)$$

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 $0 \le \epsilon < 1$.

Definition (cf. [24]): Given $0 \le \epsilon < 1$, $E \in \mathbb{R}_+$ 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.

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 an arbitrary wireless network 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 [11, Th. 1].

As before, consider a network with the node set $\{1, 2, \ldots, k\}$ where node 1 is the source node and $\mathcal{R} \subset \{2, \ldots, k\}$ 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., $u \subset \{1, 2, \ldots, k\}$ such that:

1) $1 \in u$;
2) $\mathcal{R} \not\subset u$.

Clearly, a cut u partitions the node set into two subsets: u and $u^c \triangleq \{1, \ldots, k\} \setminus u$, such that u contains the source and u^c contains at least one destination.

Let the collection of all such cuts be $U = \{u_1, u_2, \ldots, u_{|U|}\}$. Note that there are 2^{k-1} ways of assigning the nonsource nodes to either u or u^c. Out of them, 2^{k-r-1} assignments violate the second condition of a cut. Therefore, the total number of cuts is

$$|U| = 2^{k-r-1}(2^r - 1). \quad (6)$$

Given the collection of cuts U, we can construct a matrix $\mathbf{L}$ of size $|U| \times k$ as follows. The (i, j) coefficient of $\mathbf{L}$ is defined to be

$$\ell_{ij} = \begin{cases} \sum_{\ell \in u_i^c} g_{j\ell}, & \text{if } j \in u_i \\ 0, & \text{if } j \notin u_i. \end{cases} \quad (7)$$

Effectively, ℓ_{ij} is the maximum total power (in W) of the received signal at all the nodes in the destination side of a cut (i.e., set u_i^c), due to 1 W of transmission power from node $j \in u_i$. Consider the point-to-point scenario where node j is the transmitter and the destination node is a multiple-antenna system with access to received signals at all the nodes in u_i^c. Thus, this system is a single-input–multiple-output (SIMO) system with $|u_i^c|$ receiver antenna. For such a system, ℓ_{ij} is the gain G dictating the minimum energy requirement through (1) (for complete treatment of minimum energy per bit of SIMO systems, refer to [23]). Therefore, $(N_0 \log_e 2)\ell_{ij}^{-1}$ is a lower bound on the minimum energy per bit required in the simplified system where node j is a transmitter even when many nodes cooperate with no energy expenditure to decode a message. In a network setting, there could be many transmitters and several nodes could cooperate to decode the message; however, the basic idea of simplifying and reducing the network into several SIMO systems can still be used to obtain a lower bound as shown in Theorem 1.

In the following discussion, we use $\mathbf{1}_\ell$ and $\mathbf{0}_\ell$ to denote column vectors of size ℓ containing all 1s and all 0s entries, respectively.

Theorem 1: The minimum energy per bit for multicasting is lower bounded by the value E_1 of the linear program (8), whenever channel state information is not available at the transmitters

$$E_1 \triangleq \min_{\mathbf{q}\in\mathbb{R}^k} \mathbf{1}_k^T\mathbf{q} : \quad \mathbf{L}\mathbf{q} \geq (N_0 \log_e 2)\mathbf{1}_{|U|} \quad \mathbf{q} \geq \mathbf{0}_k. \tag{8}$$

To prove Theorem 1, we need Lemma 1, 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 ith element of $\mathbf{y}_u$ is given by

$$(y_u)_i = \begin{cases} y_i, & \text{if } i \in u \\ 0, & \text{if } i \notin u. \end{cases} \tag{9}$$

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, respectively. Let $\mathbf{H}$ be the random matrix formed by the fading coefficients h_{ij}, for $i,j = 1,\dots,k$. The (i,j)th entry of the fading random matrix $\mathbf{H}_u$ is given by

$$(h_u)_{ij} = \begin{cases} h_{ij}, & \text{if } j \in u \\ 0, & \text{otherwise.} \end{cases} \tag{10}$$

Dropping the time indices, we can rewrite (2) to give the observation vector at the set of nodes u^c as

$$\mathbf{y}_{u^c} = \mathbf{H}_{u^c}^T\mathbf{x} + \mathbf{z}_{u^c} \tag{11}$$

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

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

$$E_2 \triangleq \inf_{\substack{p_1,\dots,p_k\geq 0:\\ \sum_{i=1}^k p_i>0}} \max_{u\in U} \frac{\sum_{i=1}^k p_i}{\sup_{\substack{P_{\mathbf{x}}:\\ \mathbb{E}[|x_i|^2]\leq p_i\\ for\ i=1,\dots,k}} I(\mathbf{x}_u;\mathbf{y}_{u^c}|\mathbf{x}_{u^c},\mathbf{H})} \tag{12}$$

whenever channel state information is not available at the transmitters.

Proof: Appendix I. □

Proof of Theorem 1: It is sufficient to show that $E_1 \leq E_2$. To do so, we first need to simplify (12) 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 (12) by

$$I(\mathbf{x}_u;\mathbf{H}_{u^c}^T\mathbf{x}+\mathbf{z}_{u^c}|\mathbf{x}_{u^c},\mathbf{H}) = I(\mathbf{x}_u;\mathbf{H}_{u^c}^T\mathbf{x}_u+\mathbf{z}_{u^c}|\mathbf{x}_{u^c},\mathbf{H})$$
$$\leq \mathbb{E}\left[\log_2\det\left(\mathbf{I}+\frac{1}{N_0}\mathrm{cov}(\mathbf{H}_{u^c}^T\mathbf{x}_u|\mathbf{x}_{u^c},\mathbf{H})\right)\right] \tag{13}$$
$$\leq \frac{\log_2 e}{N_0}\mathbb{E}\left[\mathrm{tr}\left(\mathrm{cov}(\mathbf{H}_{u^c}^T\mathbf{x}_u|\mathbf{x}_{u^c},\mathbf{H})\right)\right] \tag{14}$$
$$= \frac{\log_2 e}{N_0}\sum_{j\in u^c}\mathbb{E}\left[\left|\sum_{i\in u} h_{ij}x_i\right|^2\right] \tag{15}$$
$$\leq \frac{\log_2 e}{N_0}\sum_{i\in u}\sum_{j\in u^c} g_{ij}p_i \tag{16}$$

where (13) is an upper bound on the mutual information under AWGN; (14) is due to Hadamard's inequality and the fact $\log(1+x) \leq x$; and (16) is obtained by maximizing the right-hand side of (15) 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 (16) in (12), we get $E_2 \geq E_3$, where

$$E_3 \triangleq \inf_{\substack{p_1,p_2,\dots,p_k\geq 0\\ :\sum_{i=1}^k p_i>0}} \left(\sum_{i=1}^k p_i\right)\frac{N_0\log_e 2}{\min_{u\in U}\sum_{i\in u}\sum_{j\in u^c} g_{ij}p_i}. \tag{17}$$

Note that E_3 is a lower bound on the minimum energy per bit.

Note from (17) 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 (17). Hence, the domain of $\mathbf{p}$ can be constrained to lie within a compact set. Therefore, the infimum in (17) can be replaced with a minimum. Suppose the optimum power policy in (17) that achieves E_3 is $\mathbf{p}^* = (p_1^*,\dots,p_k^*)$. Define

$$u^* \triangleq \mathrm{argmin}_{u\in U}\sum_{i\in u}\sum_{j\in u^c} g_{ij}p_i^*. \tag{18}$$

Next, set

$$\gamma \triangleq \frac{N_0\log_e 2}{\sum_{i\in u^*}\sum_{j\in u^{*c}} g_{ij}p_i^*} > 0. \tag{19}$$

Therefore, for all $u \in U$

$$\sum_{i\in u}\sum_{j\in u^c} g_{ij}(\gamma p_i^*) \geq N_0\log_e 2. \tag{20}$$

Therefore, all the constraints in (8) are satisfied by $\gamma\mathbf{p}^*$. Also, the values of the objective functions in (8) and (17) coincide at $\gamma\mathbf{p}^*$. Therefore, for every value of the objective function (17), the objective function (8) 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. The energy per bit allocation $\mathbf{q} = (q_1,q_2,\dots,q_k)^T \in \mathbb{R}_+^k$ is the energy of transmission (per information bit) at nodes $1,2,\dots,k$, respectively. Let $\mathcal{Q}$ be a collection of all such allocations such that the total energy per bit received at the nodes in u^c due to transmissions from nodes in u according to the allocation is at least $N_0\log_e 2$, for every cut u in the network. Any communication scheme which reliably multicasts the source message to the destination nodes should necessarily have its energy consumption per bit at its nodes in the set $\mathcal{Q}$. Thus, the particular allocation vector $\mathbf{q}^*$ in $\mathcal{Q}$ which minimizes the sum $q_1+q_2+\cdots+q_k$ is also a lower bound on the minimum total energy consumption in the network. Note that the converse may not be true, i.e., an allocation $\mathbf{q} \in \mathcal{Q}$ may not be "feasible" in the sense that no reliable communication scheme may exist which has the energy consumption per bit $\mathbf{q}$ at its nodes. Similarly, the sum-minimizing allocation $\mathbf{q}^*$ is not guaranteed to be a feasible allocation. However, in Section IV, we present some networks where decode-and-forward has energy consumption at the nodes either equal or close to $\mathbf{q}^*$.

Remark 2: In the proof of Theorem 1, (17) forms another lower bound on the minimum energy per bit which is equivalent to (8). This is stated as Corollary 1. The formulation in Corollary 1 will be useful in Section IV-B.

Corollary 1: The value E_1 of the linear program (8) satisfies $E_1 = E_3$, where

$$E_3 = \min_{\substack{\alpha_1,\alpha_2,\ldots,\alpha_k \geq 0:\\ \sum_{i=1}^k \alpha_i = 1}} \frac{N_0 \log_e 2}{\min_{u\in U} \sum_{i\in u}\sum_{j\in u^c} g_{ij}\alpha_i}. \tag{21}$$

Proof: Note that the right-hand side of (21) is the same as (17) and is obtained by replacing $p_i / \sum p_i$ with α_i in (17). Since, in Theorem 1, we have already established that $E_1 \leq E_3$, we are only left to show that $E_1 \geq E_3$ [with E_3 as given in (17)].

Since each row of $\mathbf{L}$ corresponds to a cut in U, the first constraint of (8) implies that

$$\frac{N_0 \log_e 2}{\min_{u\in U} \sum_{i\in u}\sum_{j\in u^c} g_{ij} q_i} \leq 1 \tag{22}$$

for all $\mathbf{q}$ satisfying the constraints of (8). Therefore, for all $\mathbf{q}$ satisfying the constraints of (8), the value of the linear program (8) is greater than the minima in (17), i.e.,

$$\begin{aligned}&\inf_{\substack{q_1,\ldots,q_k\geq 0:\\ \mathbf{L}\mathbf{q}\geq (N_0\log_e 2)\mathbf{1}_{|U|},\\ \sum_{i=1}^k q_i > 0}} \sum_{i=1}^k q_i \\ &\geq \inf_{\substack{p_1,\ldots,p_k\geq 0:\\ \mathbf{L}\mathbf{p}\geq (N_0\log_e 2)\mathbf{1}_{|U|},\\ \sum_{i=1}^k p_i > 0}} \left(\sum_{i=1}^k p_i\right) \frac{N_0 \log_e 2}{\min_{u\in U} \sum_{i\in u}\sum_{j\in u^c} g_{ij} p_i}.\end{aligned} \tag{23}$$

Note that since all channel gains are positive and finite, $\mathbf{q} = \mathbf{0}_k$ cannot be a solution to (8). Thus, the constraints on the left-hand side of the inequality in (23) are identical to the constraints in (8). Thus, the left-hand side of (23) has the value E_1. We can further lower bound the right-hand most side of (23) by E_3, by removing the constraint $\mathbf{L}\mathbf{p} \geq (N_0 \log_e 2)\mathbf{1}_{|U|}$. Thus, $E_3 \leq E_1$ which completes the proof. □

Remark 3: In Corollary 2, we establish another result based on the quantity 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 loss $1/G(\mathcal{S})$ for a cluster set $\mathcal{S}$ by the relation

$$G(\mathcal{S}) \triangleq \frac{1}{|\mathcal{S}|} \max_{i\in\{1,\ldots,k\}} \sum_{S\in\mathcal{S}\setminus i} g_i(S) \tag{24}$$

where

$$g_i(S) \triangleq \sum_{\ell\in S} g_{i\ell} \tag{25}$$

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. Note that the quantity $G(\mathcal{S})$ is dependent only upon the channel gains in the network. Thus, unlike program (8), evaluation of $G(\mathcal{S})$ does not require an explicit determination of optimal energy per bit allocation, which is especially convenient in analyzing the minimum energy per bit of a variety of networks, e.g., regular networks (Section IV-E). Furthermore, as shown in Corollary 2, the effective network loss $1/G(\mathcal{S})$ for networks is naturally analogous to the channel loss for point-to-point channels (1). We note that the notion of effective network loss was also treated in [11]; however, the definition in (24) leads to tighter bounds as is discussed in the proceeding Remark 4.

Corollary 2: The minimum energy per bit required to transmit a message reliably to all the nodes in $\mathcal{R}$ satisfies

$$\frac{E_b}{N_0}_{\min} \geq \frac{N_0 \log_e 2}{G(\mathcal{S})} \tag{26}$$

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

Proof: Appendix II. □

Remark 4: In [11, Th. 1], a lower bound on minimum energy per bit is derived which is the same as (26) with the specific choice of cluster set where each cluster consists of a single node from $\mathcal{R}$ and the clusters exhaust all the nodes in $\mathcal{R}$. In this case, the simplified expression for effective network loss $1/G$ is given by

$$G = \frac{1}{|\mathcal{R}|} \max_{i\in\{1,\ldots,k\}} \sum_{j\in\mathcal{R}\setminus\{i\}} g_{ij}. \tag{27}$$

The form for the network loss given in (27) simplifies the analysis of random networks where the nodes are distributed randomly and uniformly over the network area [11] and the channel gain falls off roughly as $r^{-\alpha}$ with distance $r > 0$ between nodes (see Section IV-E for more details on the path loss model). In [11], it is shown that for "dense" random networks with k nodes, $G = O\left(A_k^{-1}\right)$ when the area $A_k = o(k/\log k)$ as $k \to \infty$. On the other hand, when the node density $\lambda = k/A_k$ (in nodes/m^2) is a constant ("extended" random networks), $G = O\left(k^{-1}(\log k)^{-\alpha/2}\right)$. Furthermore, another class of grid-like networks—*regular networks*—was also considered in [11] where the network area is partitioned into square cells such that each cell holds exactly one node confined to an area in the center of the cell. As opposed to random networks, regular networks have enough structure so as to preclude the asymptotic (in k) analysis; i.e., bounds can be given for regular networks which hold even for finite k. In particular, for regular networks the gap between the upper and lower bounds on minimum energy per bit is at most a constant for all $k \geq 2$ [11, Th. 4].

However, while the lower bounds proved in [11] rely completely on the simple definition (27) of network loss, it is not possible to use the same definition for proving results for a more

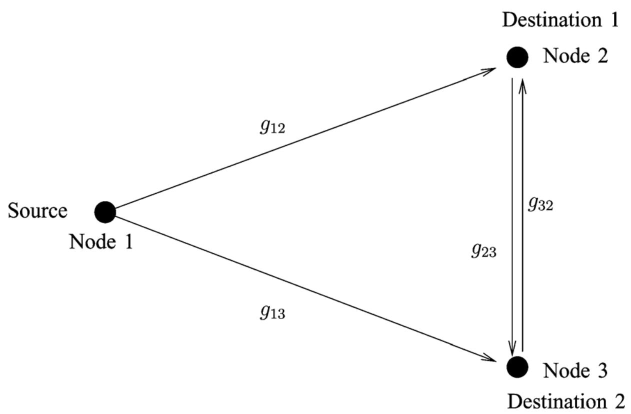

Fig. 1. Three-terminal network.

general class of regular networks considered in Section IV-E. As opposed to the definition used in [11, Sec. VI], in Section IV-E, we use a more generic definition of regular networks where there are no restrictions on the placement of nodes within the cells. This allows the nodes to be arbitrarily close to each other and thus have a maximal channel gain [say, $\Theta(1)$] between them. In the situation where nodes in groups are positioned very close to each other (thereby implying a maximum gain between them), definition (27) will give $G = O(1)$ even as the network grows sparse. Thus, for such networks, definition (27) will not give the right scaling function for G. In order to generate a tighter scaling law for G, we need to treat clusters of proximate nodes as single nodes. This can be accomplished by using the richer definition (24) of G. Indeed, as we show in Section IV-E using Corollary 2, the gap between the upper and lower bounds is still a constant even under the new, less restrictive definition of regular networks.

IV. Decode-and-Forward

In this section, we briefly describe decode-and-forward in wideband and then compare the performance of decode-and-forward with the lower bounds of the previous section for different classes of wireless networks.

A. Decode-and-Forward in the Wideband Regime

By decode-and-forward we mean the class of schemes in which the nodes completely and reliably decode the source message before transmitting anything. In the discussion on decode-and-forward, it is assumed that the receivers are noncoherent.

In [23], it was shown that in Gaussian point-to-point energy-constrained channels, achieving minimum energy per bit requires 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 receivers listen to transmissions over all the bands. Wideband broadcasts are not affected by interference from other transmissions. It was shown in [23] 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 when the accumulated energy per bit at the receiver exceeds $N_0 \log_e 2$ regardless of the knowledge of channel states of the receiver.

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

B. Three-Terminal Network

In this section, we focus on a three-terminal wireless network with one source and two cooperating destination nodes (Fig. 1). We obtain lower bounds for this setup using Corollary 1. Furthermore, we compare the performance of decode-and-forward with the lower bounds.

Using Corollary 1, we get a lower bound on the minimum energy per bit

$$E_3 = \frac{N_0 \log_e 2}{G(g_{12}, g_{13}, g_{23}, g_{32})} \tag{28}$$

where

$$G(g_{12}, g_{13}, g_{23}, g_{32}) = \max_{\substack{\alpha_1, \alpha_2, \alpha_3 \geq 0: \\ \alpha_1+\alpha_2+\alpha_3=1}} \min \Big\{ (g_{12}+g_{13})\alpha_1, g_{12}\alpha_1 + g_{32}\alpha_3, g_{13}\alpha_1 + g_{23}\alpha_2 \Big\}. \tag{29}$$

Note that the cuts used in obtaining (29) are $\{1\}$, $\{1,2\}$, and $\{1,3\}$.

For all cases of g_{12}, g_{13}, g_{23}, and g_{32}, we evaluate (28) using (29). For each of these cases, we also propose a decode-and-forward scheme and compare its performance with the lower bounds. Details on the calculations are relegated to Appendix III.

The analysis is divided into the following cases of channel gains. For each case, we assume, without loss of generality

$$g_{13} \leq g_{12}. \tag{30}$$

1) $g_{12} \leq g_{13} + g_{32}$

The three subcases are as follows.

a) $g_{23} \leq g_{13}$: In this case, $G = g_{13}$.[1] Hence

$$\frac{E_b}{N_0}_{\min} \geq \frac{\log_e 2}{g_{13}}. \tag{31}$$

The minimum energy per bit in (31) is achieved by broadcasting with enough energy per bit $[E_b = (N_0 \log_e 2/g_{13}) + \epsilon$ for any $\epsilon > 0]$ for node 3 to be able to decode the message, i.e., the total received energy per bit at node 3 exceeds $N_0 \log_e 2$. This energy is also enough for node 2 to be able to decode the message.

b) $g_{13}g_{32} \leq g_{32}g_{23} \leq g_{12}g_{23} + g_{32}g_{13}$: In this case, $G = (g_{12}g_{23})/(g_{12}+g_{23}-g_{13})$. Hence

$$\frac{E_b}{N_0}_{\min} \geq \left(\frac{1}{g_{23}} + \frac{1}{g_{12}} - \frac{g_{13}}{g_{12}g_{23}}\right) \log_e 2. \tag{32}$$

[1]In this section and Appendix III, we use the short form G to denote $G(g_{12}, g_{13}, g_{23}, g_{32})$.

This minimum energy per bit is achieved by decode-and-forward where the message is first decoded by node 2 which then forwards it to node 3. Node 3 is able to decode the message based on the initial transmission by the source node (with transmission energy per bit $E_b = (N_0 \log_e 2/g_{12}) + \epsilon$ for any $\epsilon > 0$) and the subsequent transmission by node 2 [with $E_b = (N_0 \log_e 2\,(1-(g_{13}/g_{12}))/g_{23}) + \epsilon$ for any $\epsilon > 0$].

c) $g_{12}g_{23} + g_{32}g_{13} \leq g_{32}g_{23}$: This case is different from the previous cases in that the prescribed decode-and-forward schemes do not achieve the minimum energy per bit lower bound. Gains are now in the regime where it might be better for nodes to cooperatively decode their messages rather than decode using only the direct transmission(s) from the source node. For this case

$$G = \frac{g_{12}+g_{13}}{1+\frac{g_{12}}{g_{23}}+\frac{g_{13}}{g_{32}}}. \tag{33}$$

Hence, defining

$$E_{b\,\text{lower}} = (N_0 \log_e 2)\, G^{-1} \tag{34}$$

we immediately conclude

$$E_{b\,\min} \geq E_{b\,\text{lower}} \tag{35}$$

from Corollary 2.

Note that as $g_{23}, g_{32} \to \infty$, $G \to g_{12}+g_{13}$. This is the regime where the links between the two destination nodes are so strong that the cooperation entails negligible energy, and so

$$\frac{E_b}{N_0}_{\min} \geq \frac{\log_e 2}{g_{12}+g_{13}} \tag{36}$$

which is also the minimum energy per bit of a point-to-point channel with two antennas (nodes 2 and 3) at the destination.

In general, decode-and-forward does not achieve the lower bound in (35). However, consider decode-and-forward by node 2, which has an achievable energy per bit

$$\frac{E_b}{N_0}_{DF} = \left(\frac{1}{g_{23}}+\frac{1}{g_{12}}-\frac{g_{13}}{g_{12}g_{23}}\right)\log_e 2 \tag{37}$$

as discussed in subcase b). From (34) and (37), the ratio of the upper and lower bounds is

$$\frac{E_{b\,DF}}{E_{b\,\min}} \leq \frac{g_{12}+g_{13}}{1+\frac{g_{12}}{g_{23}}+\frac{g_{13}}{g_{32}}}\,\frac{g_{12}+g_{23}-g_{13}}{g_{12}g_{23}} \tag{38}$$

$$\leq \frac{g_{12}+g_{13}}{1+\frac{g_{12}}{g_{23}}}\,\frac{g_{12}+g_{23}}{g_{12}g_{23}}$$

$$\leq \left(1+\frac{g_{13}}{g_{12}}\right) \tag{39}$$

$$\leq 2 \tag{40}$$

where the last step is due to (30). The ratio of 2 is approached when $g_{12} = g_{13} = a$, $g_{23} = g_{32} = b$, and $\frac{a}{b} \to 0$.

2) $g_{13} + g_{32} \leq g_{12}$

The two possibilities here exhibit the same behavior as the first two subcases of 1).

a) $g_{23} \leq g_{13}$: In this case, $G = g_{13}$ which implies that

$$\frac{E_b}{N_0}_{\min} \geq \frac{\log_e 2}{g_{13}} \tag{41}$$

The minimum energy per bit is achieved by a broadcast with enough energy to reach both nodes 2 and 3.

b) $g_{13} \leq g_{23}$: In this case, $G = (g_{12}g_{23})/(g_{12}+g_{23}-g_{13})$. Hence

$$\frac{E_b}{N_0}_{\min} \geq \left(\frac{1}{g_{23}}+\frac{1}{g_{12}}-\frac{g_{13}}{g_{12}g_{23}}\right)\log_e 2. \tag{42}$$

This minimum energy per bit is achieved by decode-and-forward from node 2 to node 3 as discussed in case 1), subcase b).

Remark 5: When decode-and-forward is suboptimal, an estimate-and-forward-based scheme [6] has been proposed in [12]. The idea involves node 2 to quantize its observation at a rate high enough so that node 3 could decode the source message given its own observation and the digital message from node 2. Once node 3 has decoded the message successfully, it can make a suitable transmission to node 2 so that node 2 could also decode the source message now. As shown in [12], this scheme reduces the maximum gap between the upper and lower bounds from a factor of 2 to a factor of 1.74 for symmetric channel gains. In fact, as $g_{23} = g_{32}$ approach infinity, the gap reduces to 1.

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 $t = 2, \ldots, k$, update
   - $S(t) = S(t-1) \bigcup \{j(t-1)\}$
   - $j(t) = \text{argmax}_{j\in\{1,\ldots,k\}\setminus S(t)} \sum_{\ell\in S(t)} \tilde{q}_\ell\, g_{\ell j}$
   - $\tilde{q}_{j(t)} = (k-t)\, q_{j(t)}$

Fig. 2. The algorithm to determine parameters $j(t)$ and $\tilde{q}_{j(t)}$.

C. Performance of Decode-and-Forward in Arbitrary Networks

For three-terminal networks, we have seen that decode-and-forward is suboptimal by at most 3 dB. In this section, we generalize this result for arbitrary networks.

Theorem 2: In an arbitrary network with k nodes, decode-and-forward requires energy per bit $E_{b\,DF}$ bounded by

$$E_{b\,DF} \leq (k-1) E_{b\,\min}. \tag{43}$$

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

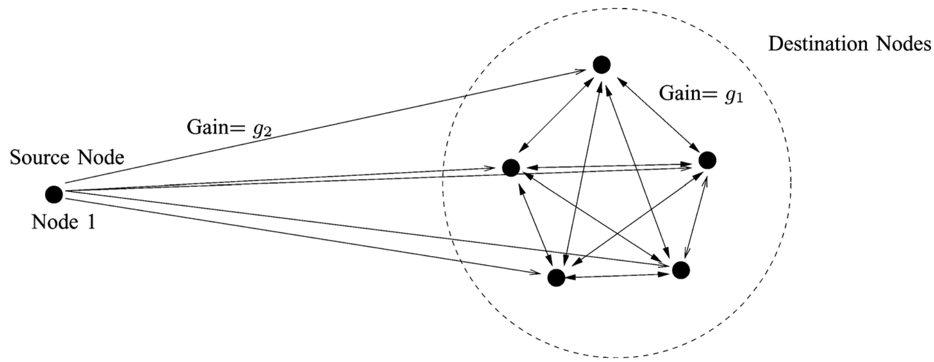

Fig. 3. A network with isolated source and five destinations.

Suppose that the energy policy $\mathbf{q} = (q_1, \ldots, q_k)$ achieves the minimum value of the linear program (8) for the given network, i.e.,

$$\frac{E_b}{N_0}_{\min} \geq \sum_{i=1}^{k} q_i. \tag{44}$$

Using $\mathbf{q}$, 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 that has already decoded the message. The transmission power levels at the nodes is determined by the policy $\tilde{\mathbf{q}}$.

Choose any $\epsilon > 0$. The proposed decode-and-forward scheme is the following. Node $j(t)$ transmits the message (in wideband) with energy expenditure per bit $\tilde{q}_{j(t)} + \epsilon$ in time slot t, for $t = 1, \ldots, k-1$. If node $j(t)$ is unable to decode the message by the end of time slot $t-1$, an error is declared and the algorithm terminates. Parameters $j(t)$ and $\tilde{q}_{j(t)}$ are recursively defined by Algorithm 1 in Fig. 2. Note that $\tilde{q}_{j(k)} = 0$, so no transmission is required in the slot k.

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 Lemma 2 and (44).

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

1) the set $\{j(1), j(2), \ldots, j(k)\}$ forms a permutation of $\{1, 2, \ldots, k\}$;
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 $t = 2, \ldots, k$.

Proof: Appendix IV. □

Remark 6: When $k = 2$, i.e., for point-to-point channels, Theorem 2 gives tight results [23], i.e., $E_{b\mathrm{DF}} = E_{b\min}$. Of course, decode-and-forward is trivial in this case.

Remark 7: The result of Theorem 2 is consistent with the result for the three-terminal network considered in Section IV-B (see also [12] for a more extensive treatment) 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.

In order to further understand the tightness of the result in Theorem 2, consider a network where there is an isolated source node and $k-1$ destination nodes which are close to each other but far away from the source node. In Fig. 3, we show the isolated node which has a gain of g_2 to each of the $k-1$ ($= 5$ in Fig. 3) destination nodes, and all the destination nodes have a gain of g_1 to each other. We focus on the particular case of when g_1 is much larger than g_2.

For such a network, the energy requirement of decode-and-forward is

$$E_{b\mathrm{DF}} = \frac{N_0 \log_e 2}{g_2} \tag{45}$$

at the source node, since if none of the destination nodes are allowed to transmit anything before they decode the message, the initial transmission from the source should be powerful enough to let at least one of the destinations decode the message. Thus, $E_{b\mathrm{DF}}$ in (45) is the minimum energy requirement of decode-and-forward; furthermore, $E_{b\mathrm{DF}}$ is also sufficient for all the destinations to be able to decode the message reliably. On the other hand, consider the energy per bit quota

$$\mathbf{q}_1 = \left(\frac{N_0 \log_e 2}{(k-1)g_2}, \frac{N_0 \log_e 2}{(k-1)g_1}, \ldots, \frac{N_0 \log_e 2}{(k-1)g_1} \right)^T \tag{46}$$

in the context of linear program (8), where the source node has a minimum energy per bit allocation of $N_0 \log_e 2/(k-1)g_2$ and the rest of the nodes have an allocation of $N_0 \log_e 2/(k-1)g_1$. The allocation vector $\mathbf{q}_1$ satisfies all the constraints of program (8), and thus is a feasible allocation. Hence, according to Theorem 1, the lower bound E_1 on the minimum energy per bit satisfies

$$E_1 \leq \left(\frac{1}{(k-1)g_2} + \frac{1}{g_1} \right) N_0 \log_e 2 \tag{47}$$

which immediately implies that the gap between the upper bound $E_{b\mathrm{DF}}$ and the lower bound E_1 from Theorem 1 approaches a factor of $(k-1)$ as g_1 becomes large in comparison with $(k-1)g_2$. Thus, the network described in Fig. 3 is an example of a network where it is not possible to further improve upon Theorem 2 using the results presented in this paper. In order to obtain tighter results, we would need to either improve the lower bounds or develop more efficient communication schemes.

Remark 8: The scheme obtained from Algorithm 1 need not be the best possible decode-and-forward scheme. Optimal decode-and-forward schemes are proposed in [10], [19], and [20]. As mentioned in Section I, the optimal decode-and-forward schemes discussed in [10], [19], and [20] require first determining an optimal transmission order of nodes, and then optimizing the energy allocation for this order. However, figuring out the minimum energy consumption of decode-and-forward using this method is nontrivial for the most general cases of networks. Instead, the decode-and-forward scheme described in Algorithm 1 is more suitable for our analysis.

D. Decode-and-Forward for DAGs

Consider 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 DAG $\mathcal{G}(V, E)$ 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

$$(i, j) \notin E \iff g_{ij} = 0 \tag{48}$$

i.e., there is no path from a node to itself that is over links with strictly positive channel gains. For a meaningful discussion, we assume that for every nonsource node there is a path that takes us from the source node to it. This also implies that no edge points into the source. The result for this setup is given in Theorem 3.

Note that examples of networks which can be represented as DAGs include relay networks and broadcast channels. In many resource constrained wireless networks, nodes may have the ability to transmit and receive only over selected frequency bands which may differ from node to node. This could introduce directionality into the communication system due to different channel gains on different links. Furthermore, some nodes may not be able to receive messages in certain frequency bands while they could broadcast in other (more universal) bands.

Theorem 3: For a wireless network represented by a DAG $\mathcal{G}$, the minimum energy per bit $E_{b\min}$ for broadcasting is given by the value E_1 of the linear program (8).

Proof: The converse part is directly from Theorem 1. Let the minimum value of the linear program (8) 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., [13, Ch. 2.2.3] and [5, Ch. 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. The next result shows that the probability of error is arbitrarily small.

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

Proof: Appendix V. □

Note that since the choice of $\epsilon > 0$ is arbitrary, we can have an achievable energy per bit equal to $\sum_{i=1}^{k} q_i$. This immediately implies that the value of program (8) is achievable. Note that node k need not transmit anything since there are no outgoing edges from node k rendering its transmissions useless for any other node. □

E. Regular Networks

In this section, we study energy efficiency in *regular networks* which were also considered in [11] in a less general setup. The bounds of [11] are not tight enough to deal with the general placement of nodes. Also, here, we allow more than one node per cell. The deterministic approach taken here is similar to the approach taken before in [14].

In a regular network, $k \ge 2$ 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:

1) at most $\bar{k}$ nonsource nodes;
2) at least $\underline{k} \ge 1$ nonsource 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 particular 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

$$\frac{k}{\bar{k}} \le n^2 \le \frac{k}{\underline{k}}. \tag{49}$$

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

$$g(r) = r^{-\alpha} \tag{50}$$

where $r_0 > 0$ and $\alpha > 2$ are constants of the model. We impose an additional restriction on the path loss function that for some constant $\bar{g} > 0$

$$g(r) \le \bar{g}, \qquad \text{for all } r \ge 0 \tag{51}$$

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 2. 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. 4). We operate under the *minimal information* framework of [11] which requires the nonsource nodes to make

$\texttt{FLOOD}(E_{b_1}, E_{b_2})$ [11]

1) The source node transmits only in the 1^{st} time slot with energy per bit E_{b1}.
2) At the beginning of time slot $t = 2, ..., T$, 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_{b2}.
   - Else, keep quiet.

Fig. 4. $\texttt{FLOOD}(E_{b_1}, E_{b_2})$.

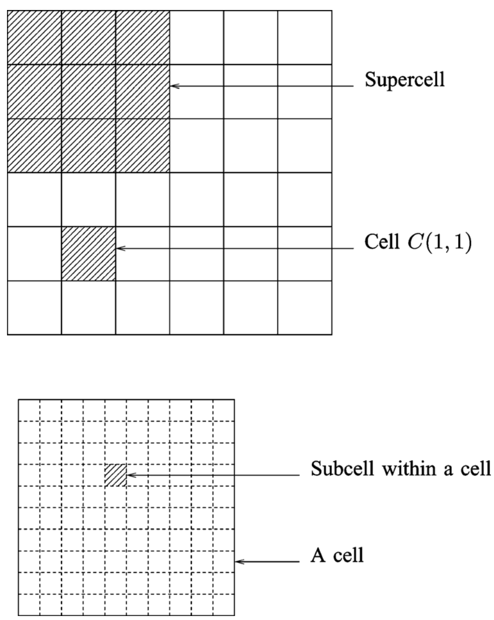

Fig. 5. A regular network with 36 cells and four supercells each consisting of nine cells. Also shown is a single cell with 100 subcells.

transmissions at preallocated energy levels. In other words, the nodes do not need to know their locations for the flooding to be successful since the location information does not determine any parameter in the FLOOD algorithm as described in Fig. 4. Furthermore, we also assume that the channel state information is not available at the receivers. The result for regular networks is given in Theorem 4.

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

$$\frac{E_{b\text{flood}}}{E_{b\min}} \leq c\,\frac{\bar{k}^{\alpha+2}}{\underline{k}} \tag{52}$$

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

Proof: We begin with 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 nine cells each. Each supercell consists of a *leader cell* of the form $C(3m_1+1, 3m_2+1)$ for $m_1, m_2 = 0, 1, \ldots, (n/3)-1$, and the eight cells (*subordinate cells*) that are tangent to the leader cell. Note that there are $n^2/9$ supercells in the network. (See Fig. 5.)

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

1) Initialize

$$S_{m_1,m_2} = \{\text{All nodes in the cell } C(3m_1+1, 3m_2+1)\}$$

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

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. 6. Applying Algorithm 2 to each leader cell gives us a cluster set $\mathcal{S} = \{S_{m_1,m_2} : m_1, m_2 \in \{0, \ldots, (n/3)-1\}\}$. An example of a cluster set constructed according to Algorithm 2 is shown in Fig. 7. A few observations about the clusters and the cluster set are given in the following result.

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

1) For any cluster in $\mathcal{S}$, the minimum distance between a node inside the cluster and a node outside the cluster is at least $s/(9\bar{k}+1)$.
2) 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) $|\mathcal{S}| = n^2/9$.
4) Each cluster has at least $\underline{k}$ nodes and at most $9\bar{k}$.

Proof: Appendix VI. □

Knowing $\mathcal{S}$, the converse is obtained from Corollary 2. In order to apply Corollary 2, we need to upper bound the following quantity for all $i = 1, \ldots, k$:

$$\sum_{S \in \mathcal{S} \setminus i} g_i(S) \tag{53}$$

which is the sum of gains from node i to all nodes that 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, quantity in (53) can instead 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 $\ell = 0, 1, \ldots, (n/3)-1$. 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 (24) and monotonicity of $g(\cdot)$, we obtain

$$G(\mathcal{S}) \leq \frac{1}{|\mathcal{S}|}(9\bar{k})\left(9g\left(\frac{s}{(9\bar{k}+1)}\right) + \sum_{\ell=2}^{n/3-1} 8\ell\, g\left(3(\ell-1)s\right)\right) \tag{54}$$

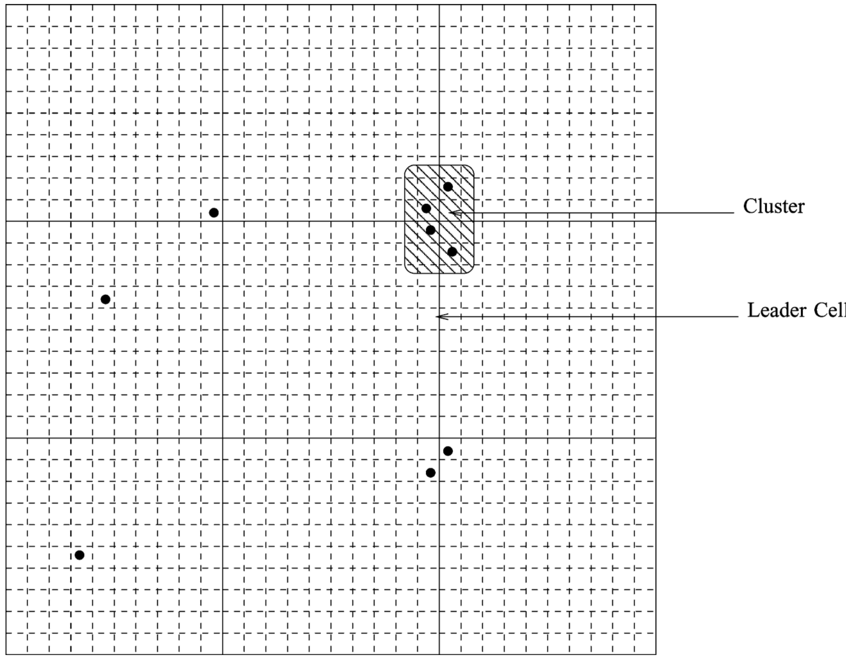

Fig. 7. An example of a cluster set, shown for a single supercell and $\bar{k} = 1$.

$$\leq \frac{729\,\bar{k}^2}{k}\left(g\left(\frac{s}{(9\bar{k}+1)}\right) + \sum_{\ell=2}^{n/3-1} \ell\, g\,(3(\ell-1)s)\right). \tag{55}$$

The proof is divided into three cases.

- *Case 1:* $ns/(9\bar{k}+1) < r_0$.
  In this case, the whole network can be enclosed within a box of side $(9\bar{k}+1)r_0$. Thus, the following direct transmission scheme:
  $$\texttt{FLOOD}\left(\frac{N_0 \log_e 2}{g(\sqrt{2}(9\bar{k}+1)r_0)} + \epsilon_1, 0\right) \tag{56}$$
  for any $\epsilon_1 > 0$, is able to reliably communicate to all the nodes. Since the choice of $\epsilon_1 > 0$ is arbitrary, the achievable energy per bit of (56) is
  $$\frac{E_b}{N_0}_{\text{flood}} = \left(\sqrt{2}r_0(9\bar{k}+1)\right)^{\alpha} \log_e 2 \tag{57}$$
  where we have used (50) and the fact that the energy consumption of $\texttt{FLOOD}(E_{b1}, E_{b2})$ is at most $E_{b1}+(k-1)E_{b2}$, to obtain (57). On the other hand, since the gain to a node cannot exceed $\bar{g}$, the effective network radius (in the sense of [11, Th. 1]) does not exceed $\bar{g}$. Therefore, by [11, Th. 1]
  $$\frac{E_b}{N_0}_{\min} \geq \frac{\log_e 2}{\bar{g}}. \tag{58}$$
  Therefore, from (57) and (58)
  $$\frac{E_{b\text{flood}}}{E_{b\min}} \leq \bar{g}\left(\sqrt{2}r_0(9\bar{k}+1)\right)^{\alpha} \leq c_1\bar{k}^{\alpha} \tag{59}$$
  for some positive constant c_1.

- *Case 2:* $r_0 < s/(9\bar{k}+1)$.
  In this case, all the distances of interest in (55) are greater than r_0. Using (50), we can bound $G(\mathcal{S})$ as
  $$G(\mathcal{S}) \leq \frac{729\,\bar{k}^2}{k}\left(\frac{(9\bar{k}+1)^{\alpha}}{s^{\alpha}} + \sum_{\ell=2}^{\infty} \frac{\ell}{3^{\alpha}(\ell-1)^{\alpha}s^{\alpha}}\right) \tag{60}$$
  $$\leq c_2\frac{\bar{k}^{\alpha+2}}{ks^{\alpha}} \tag{61}$$
  for some positive constant c_2. In deriving (61), we have made use of the facts that $\ell/(\ell-1)^{\alpha} \leq 2^{\alpha} l^{1-\alpha}$ for $\ell \geq 2$, and that $\sum_{\ell=1}^{\infty} \ell^{1-\alpha} < \infty$ since $\alpha > 2$. So, from Corollary 2 and (61)
  $$\frac{E_b}{N_0}_{\min} \geq \frac{ks^{\alpha}\log_e 2}{c_2\bar{k}^{\alpha+2}}. \tag{62}$$
  Next
  $$\texttt{FLOOD}\left(\frac{N_0 \log_e 2}{g(2\sqrt{2}s)} + \epsilon_1, \frac{N_0 \log_e 2}{\underline{k}g(2\sqrt{2}s)} + \epsilon_1\right) \tag{63}$$
  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 $2\sqrt{2}s$ (i.e., two diagonals). Note that the maximum distance between nodes in adjacent cells cannot be more than $2\sqrt{2}s$. Thus, after the first transmission, if all the nodes (at least $\underline{k}$ in number) in a cell have already decoded a message, then their transmissions according to (63) 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.
  Since $2\sqrt{2}s > r_0$ and the choice of $\epsilon_1 > 0$ is arbitrary, we get the following achievable energy per bit for (63):
  $$\frac{E_b}{N_0}_{\text{flood}} \leq 8^{\alpha/2}s^{\alpha}\left(1 + \frac{k-1}{\underline{k}}\right) N_0 \log_e 2. \tag{64}$$

Therefore, for this case

$$\frac{E_{b\text{flood}}}{E_{b\min}} \leq c_3 \frac{\bar{k}^{\alpha+2}}{\underline{k}} \tag{65}$$

for some positive constant c_3.

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

$$L \triangleq \left\lfloor \frac{(9\bar{k}+1)r_0}{s} \right\rfloor \geq 1. \tag{66}$$

Since $L \geq 1$ due to the defining condition of this case, note that

$$\frac{1}{2}\frac{(9\bar{k}+1)r_0}{s} < L \leq \frac{(9\bar{k}+1)r_0}{s} \tag{67}$$

and

$$\frac{(9\bar{k}+1)r_0}{s} < L+1 \leq 2\,\frac{(9\bar{k}+1)r_0}{s} \tag{68}$$

since $(L+1)/2 \leq L$.
Next, G can be bounded as

$$\begin{aligned} G(\mathcal{S}) &\leq \frac{729\,\bar{k}^2}{k}\left(\bar{g} + \sum_{\ell=2}^{L}\ell\bar{g} + \sum_{\ell=L+1}^{n/3-1}\frac{\ell}{3^\alpha(\ell-1)^\alpha s^\alpha}\right) && (69)\\ &\leq \frac{729\,\bar{k}^2}{k}\left(\frac{L(L+1)}{2}\bar{g} + \sum_{\ell=L+1}^{\infty}\frac{\ell}{3^\alpha(\ell-1)^\alpha s^\alpha}\right) && (70)\\ &\leq \frac{729\,\bar{k}^2}{k}\left(\frac{(9\bar{k}+1)^2 r_0^2}{s^2}\bar{g} + \frac{2^\alpha}{3^\alpha(\alpha-2)}\frac{1}{s^\alpha}\frac{1}{L^{\alpha-2}}\right) && (71)\\ &\leq \frac{729\,\bar{k}^2 r_0^{2-\alpha}}{k\,s^2} \times\left((9\bar{k}+1)^2\bar{g}r_0^\alpha + \frac{2^{2\alpha-2}}{3^\alpha(\alpha-2)}\frac{1}{(9\bar{k}+1)^{\alpha-2}}\right) && (72)\\ &\leq c_4\frac{\bar{k}^4}{ks^2} && (73)\end{aligned}$$

for some positive constant c_4. Inequality (69) is directly from (55) by separating the summation from $\ell = 2$ to L, and by upperbounding the gains $g\left(s/(9\bar{k}+1)\right)$ and $g(3(\ell-1)s) \leq \bar{g}$ for all $\ell \leq L$ by $\bar{g}$; inequality (70) is by noting that $1 + \sum_{\ell=2}^{L}\ell = L(L+1)/2$ and by extending the upper limit of second summation from $n/3-1$ to ∞; the first term in (71) is obtained from inequalities (67) and (68) and the second term is obtained by noting that

$$\begin{aligned}\sum_{\ell=L+1}^{\infty}\frac{\ell}{(\ell-1)^\alpha} &\leq 2^\alpha\sum_{\ell=L+1}^{\infty}\ell^{1-\alpha} && (74)\\ &\leq 2^\alpha\int_L^\infty u^{1-\alpha}\,du && (75)\\ &= \frac{2^\alpha}{\alpha-2}L^{2-\alpha} && (76)\end{aligned}$$

where (74) is due to the fact that $\ell/2 \leq \ell-1$ for all $\ell \geq L+1$, since $L \geq 1$. Thus, (76) leads to (71). Inequality (72) is a direct consequence of (67). Finally, (73) is obtained by noting that the first term of (72) is greater than a constant times the second term [since $\bar{g}r_0^\alpha \geq 1$ due to (51)]. So, from Corollary 2, (73) gives us

$$\frac{E_b}{N_0}_{\min} \geq \frac{ks^2}{c_4\bar{k}^4}. \tag{77}$$

Furthermore

$$\texttt{FLOOD}\left(\frac{N_0\log_e 2}{g(2\sqrt{2}sL)} + \epsilon_1, \frac{N_0\log_e 2}{\underline{k}\sum_{\ell=1}^{L}\ell g(2\sqrt{2}s\ell)} + \epsilon_1\right) \tag{78}$$

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

$$\mathcal{S}_T \triangleq \{C(x,y) : \max\{x,y\} \leq (L-1)+T\} \tag{79}$$

for $1 \leq T < n-L+1$. Next, let us assume that all the nodes in the cell $\mathcal{S}_T$ decode the message by the end of time slot T. We now show that by the end of time slot $T+1$, all nodes in the set $\mathcal{S}_{T+1}$ decode the message with vanishing probability of error. It is enough to show this for $\mathcal{S}_{T+1}\setminus\mathcal{S}_T$ since, by our hypothesis, the nodes in $\mathcal{S}_T$ have already decoded the message. Without loss of generality, consider the nodes in the cell $C(L+T,Y) \in \mathcal{S}_{T+1}\setminus\mathcal{S}_T$ where $Y \in \{0,1,\ldots,L+T\}$. For every $1 \leq \ell \leq L$, the set of cells

$$\begin{aligned}\mathcal{V}_\ell \triangleq \{C(L+T-\ell,y') : |y'-Y| \leq \ell,\\ y' \in \{0,1,\ldots,L+T-1\}\}\end{aligned} \tag{80}$$

is a subset of $\mathcal{S}_T$. Furthermore, if $\ell \neq \ell'$, then $\mathcal{V}_\ell$ and $\mathcal{V}_{\ell'}$ are disjoint. Note also that each cell in $\mathcal{V}_\ell$ is at most ℓ steps away from $C(L+T,Y)$, each step being either in the vertical, horizontal, or diagonal directions. Thus, the nodes in $C(L+T,Y)$ are at most at a distance of $2\sqrt{2}s\ell$ from the nodes in $\mathcal{V}_\ell$. Thus, the total energy (per bit) collected at any node in $C(L+T,Y)$ due to transmissions from $\mathcal{V}_\ell$ according to (78) is greater than

$$\underline{k}\ell g(2\sqrt{2}s\ell)\left(\frac{N_0\log_e 2}{\underline{k}\sum_{\ell=1}^{L}\ell g(2\sqrt{2}s\ell)} + \epsilon_1\right) \tag{81}$$

since the cardinality of $\mathcal{V}_\ell$ is greater than ℓ. Aggregating the energy due to transmissions from $\mathcal{V}_\ell$ for all $\ell \in \{1,\ldots,L\}$ gives us that the total received energy per bit at any node in $C(L+T,Y)$ (and thus, at any node in $\mathcal{S}_{T+1}\setminus\mathcal{S}_T$) is greater than $N_0\log_e 2$, which makes reliable decoding possible. By induction, this argument works for all $T < n-L+1$, and hence for all nodes in the network.

Since $L \leq (9\bar{k}+1)r_0/s$, for all $\ell \leq L$

$$g(2\sqrt{2}s\ell) \geq g(2\sqrt{2}sL) \tag{82}$$

$$\geq \left(2\sqrt{2}s\frac{(9\bar{k}+1)r_0}{s}\right)^{-\alpha} \quad (83)$$

$$= 8^{-\alpha/2}\ \left((9\bar{k}+1)r_0\right)^{-\alpha}. \quad (84)$$

Using (84) and (78), the energy consumption due to non-source nodes is not more than

$$k\left(\frac{N_0(\log_e 2)8^{\alpha/2}\left((9\bar{k}+1)r_0\right)^{\alpha}}{\underline{k}\sum_{\ell=1}^{L}\ell}+\epsilon_1\right) \quad (85)$$

$$\leq k\left(\frac{2N_0(\log_e 2)8^{\alpha/2}\left((9\bar{k}+1)r_0\right)^{\alpha}}{\underline{k}L^2}+\epsilon_1\right) \quad (86)$$

$$\leq \frac{c_5 k\bar{k}^{\alpha-2}s^2}{\underline{k}} \quad (87)$$

for some constant c_5, where we have assumed an appropriate choice of $\epsilon_1 > 0$ and used (67) to obtain (87). Next, using (84), we can show that the energy consumption by source node is at most $N_0(\log_e 2)8^{\alpha/2}\left((9\bar{k}+1)r_0\right)^{\alpha}+\epsilon_1$, which is less than the right-hand side of (86) even when $L = n$ [since $n^2\underline{k} \leq k$ from (49)]. Thus, for an appropriate c_5, the right-hand side of (87) is also an upper bound on the total energy consumption per bit for *all* the nodes in the network, i.e., the total energy consumption of (78) is

$$\frac{E_b}{N_0}_{\text{flood}} \leq c_5 k\bar{k}^{\alpha-2}s^2/\underline{k}. \quad (88)$$

Therefore, from (77)–(88)

$$\frac{E_{b\text{flood}}}{E_{b\min}} \leq c_4 c_5\frac{\bar{k}^{\alpha+2}}{\underline{k}}. \quad (89)$$

Equation (52) is now immediate by combining all the cases. □

Remark 9: Suppose that there are $k-1$ nonsource nodes in a network that 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 $k \to \infty$[14, Claim 3.1]. Therefore, for random placement of nodes, by Theorem 4, we get

$$\frac{E_{b\text{flood}}}{E_{b\min}} = O\left((\log k)^{\alpha+2}\right). \quad (90)$$

V. Conclusion

We have studied the maximum possible energy efficiency of reliable multicasting in wireless networks by focusing on the minimum energy per bit.

An information-theoretic lower bound on the minimum energy per bit for multicasting in arbitrary networks is derived in terms of a linear program. For the achievable part, we have proposed a decode-and-forward-based communication scheme that operates in the wideband regime and does not require knowledge of the channel states at the receivers.

For networks with one source and two destination nodes that cooperate, decode-and-forward has been shown to achieve the minimum energy per bit under various conditions. In situations where the channel gains between the destination nodes are large, decode-and-forward is not optimal but still is worse by at most 3 dB.

For arbitrary networks with k nodes, decode-and-forward has at most $k-1$ times the energy expenditure of the lower bound. Thus, this result can be thought of as the generalization of the factor of 2 in the three-terminal network to a factor of $k-1$ in arbitrary networks. This result also implies that the performance of decode-and-forward cannot be arbitrarily worse than that of the optimal scheme. Furthermore, the maximal gap is not too large for smaller networks, and 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 DAGs. 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 retransmit it for the benefit of their peers.

Although the gap between the upper and lower bounds can be large in arbitrary networks, the situation could be different in the real-world networks. For example, if the number of nodes within each geographical cell is fixed (regular networks) the gap is at most a constant independent of the size of the network. Regular networks can also be used to show that the gap is at most a poly-logarithmic factor in the number of nodes for large random networks. Thus, regular networks demonstrate the utility of both our lower bound and simple decode-and-forward schemes in many setups of interest.

Appendix I
Proof of Lemma 1

Proof (Straightforward generalization of [11, Lemma 1]): Consider an $(n, M, E_{\text{total}}, \epsilon)$ code over n channel uses. The total energy consumption at node i is

$$E_i = \sum_{t=1}^{n} E_{i,t} \quad (91)$$

which satisfies

$$\sum_{i=1}^{k} E_i = E_{\text{total}}. \quad (92)$$

For the rest of the proof, $\mathbf{x}_u^{(n)} = (x_i^{(n)})_{i\in u}$ denotes the collection of transmissions by the nodes in set $u \subset \{1, 2, \ldots, k\}$. $\mathbf{x}_{u,t} = (x_{i,t})_{i\in u}$ is the set of symbols transmitted by the nodes in set u at time t. We use $H(t)$ to denote the matrix of channel coefficients at time t, i.e., the ith, jth entry of $H(t)$ is the fading coefficient $h_{ij}(t)$. Furthermore, the message random variable is denoted by m.

Fix a cut $u \in U$. By the definition of u, we have $1 \in u$ and a destination node $j \in u^c$. For the cut u, we derive a form of *cut-set* or *max-flow min-cut* bound (see [6, Th. 4] and [7, Th. 15.10.1]) in the following steps:

$$(1-\epsilon)\log_2 M$$
$$\leq I(m; y_j^{(n)}) + 1 \quad (93)$$

$$\leq I(m;\mathbf{y}_{u^c}^{(n)})+1 \quad (94)$$

$$=\sum_{t=1}^{n} I(m;\mathbf{y}_{u^c,t}|\mathbf{y}_{u^c,1},\ldots,\mathbf{y}_{u^c,t-1})+1 \quad (95)$$

$$\leq\sum_{t=1}^{n} I(m;\mathbf{y}_{u^c,t},\mathbf{x}_{u^c,t}|\mathbf{y}_{u^c,1},\ldots,\mathbf{y}_{u^c,t-1})+1 \quad (96)$$

$$=\sum_{t=1}^{n} I(m;\mathbf{y}_{u^c,t}|\mathbf{y}_{u^c,1},\ldots,\mathbf{y}_{u^c,t-1},\mathbf{x}_{u^c,t})+1 \quad (97)$$

$$\leq\sum_{t=1}^{n} I(m,\mathbf{x}_{u,t};\mathbf{y}_{u^c,t}|\mathbf{y}_{u^c,1},\ldots,\mathbf{y}_{u^c,t-1},\mathbf{x}_{u^c,t})+1 \quad (98)$$

$$=\sum_{t=1}^{n}\Big(I(\mathbf{x}_{u,t};\mathbf{y}_{u^c,t}|\mathbf{y}_{u^c,1},\ldots,\mathbf{y}_{u^c,t-1},\mathbf{x}_{u^c,t}) + I(m;\mathbf{y}_{u^c,t}|\mathbf{y}_{u^c,1},\ldots,\mathbf{y}_{u^c,t-1},\mathbf{x}_{u,t},\mathbf{x}_{u^c,t})\Big)+1 \quad (99)$$

$$\leq\sum_{t=1}^{n} I(\mathbf{x}_{u,t};\mathbf{y}_{u^c,t}|\mathbf{x}_{u^c,t})+1 \quad (100)$$

$$\leq\sum_{t=1}^{n} I(\mathbf{x}_{u,t};\mathbf{y}_{u^c,t}|\mathbf{x}_{u^c,t},\mathbf{H}(t))+1 \quad (101)$$

$$\leq\sum_{t=1}^{n}\sup_{\substack{P_{\mathbf{x}_t}:\\ \mathbb{E}[|x_{i,t}|^2]\leq E_{i,t}\\ \text{for } i=1,\ldots,k}} I(\mathbf{x}_{u,t};\mathbf{y}_{u^c,t}|\mathbf{x}_{u^c,t},\mathbf{H}(t))+1 \quad (102)$$

$$\leq\sum_{t=1}^{n}\sup_{\substack{P_{\mathbf{x}_1\mathbf{x}_2\ldots\mathbf{x}_n}:\\ \frac{1}{n}\sum_{t=1}^{n}\mathbb{E}[|x_{i,t}|^2]\leq\frac{1}{n}\sum_{t=1}^{n}E_{i,t}\\ \text{for } i=1,\ldots,k}} I(\mathbf{x}_{u,t};\mathbf{y}_{u^c,t}|\mathbf{x}_{u^c,t},\mathbf{H}(t)) +1 \quad (103)$$

$$=n\sup_{\substack{P_{\mathbf{x}_1\mathbf{x}_2\ldots\mathbf{x}_n}:\\ \frac{1}{n}\sum_{t=1}^{n}\mathbb{E}[|x_{i,t}|^2]\leq\frac{1}{n}E_i\\ \text{for } i=1,\ldots,k}} \frac{1}{n}\sum_{t=1}^{n} I(\mathbf{x}_{u,t};\mathbf{y}_{u^c,t}|\mathbf{x}_{u^c,t},\mathbf{H}(t)) +1 \quad (104)$$

$$=n\sup_{\substack{P_{\mathbf{x}_q}:\\ \mathbb{E}[|x_{i,q}|^2]\leq\frac{1}{n}E_i\\ \text{for } i=1,\ldots,k}} I(\mathbf{x}_{u,q};\mathbf{y}_{u^c,q}|\mathbf{x}_{u^c,q},\mathbf{H}(q),q)+1 \quad (105)$$

$$\leq n\sup_{\substack{P_{\mathbf{x}_q}:\\ \mathbb{E}[|x_{i,q}|^2]\leq\frac{1}{n}E_i\\ \text{for } i=1,\ldots,k}} I(\mathbf{x}_{u,q};\mathbf{y}_{u^c,q}|\mathbf{x}_{u^c,q},\mathbf{H}(q))+1 \quad (106)$$

$$=n\sup_{\substack{P_{\mathbf{x}}:\\ \mathbb{E}[|x_i|^2]\leq E_i/n}} I(\mathbf{x}_u;\mathbf{y}_{u^c}|\mathbf{x}_{u^c},\mathbf{H})+1 \quad (107)$$

where (93) is due to Fano's inequality; inequality (94) is by expanding the set of random variables; applying the chain rule for mutual information to (94) gives us (95); inequality (96) is again by expanding the set of random variables; step (97) is obtained by noticing that $\mathbf{x}_{u^c,t}$ is a function of $\mathbf{y}_{u^c,1},\ldots,\mathbf{y}_{u^c,t-1}$ and hence, independent of m when conditioned over $\mathbf{y}_{u^c,1},\ldots,\mathbf{y}_{u^c,t-1}$; steps (98) and (99) are obtained by expanding the set of random variables and then applying the chain rule; step (100) is obtained by noting that if random variables $Y-X-A$ form a Markov chain then $I(X;Y|A)\leq I(X;Y)$, thus the upper bound on the first mutual information term in (99) can be obtained by noticing that, conditioned on $\mathbf{x}_{u^c,t}$, $\mathbf{y}_{u^c,t}-\mathbf{x}_{u,t}-\{\mathbf{y}_{u^c,1},\ldots,\mathbf{y}_{u^c,t-1}\}$ form a Markov chain, and that the second mutual information term in (99) is zero due to the fact that $\mathbf{y}_{u^c,t}$ is independent of m given $\mathbf{x}_{u,t}$ and $\mathbf{x}_{u^c,t}$; step (101) is a result of the fact that, conditioned on $\mathbf{x}_{u^c,t}$, the transmission $\mathbf{x}_{u,t}$ is independent of the channel state $\mathbf{H}(t)$ since channel state information is not available at the transmitters; applying the energy constraints $E_{i,t}$ on the symbols $x_{i,t}$ gives us (102), where $\mathbf{x}_t$ denotes the transmission vector at all the nodes at time t; (103) and (104) hold since we expand the set of joint distributions $P_{\mathbf{x}_1\mathbf{x}_2\ldots\mathbf{x}_n}$ over which the supremum is considered, from $P_{\mathbf{x}_1\mathbf{x}_2\ldots\mathbf{x}_n}:\mathbb{E}[|x_{i,t}|^2]\leq E_{i,t}\,\forall\, i=1,\ldots,k$ and $t=1,\ldots,n$, to $P_{\mathbf{x}_1\mathbf{x}_2\ldots\mathbf{x}_n}:\frac{1}{n}\mathbb{E}[|x_{i,t}|^2]\leq\frac{1}{n}E_{i,t}\,\forall\, i=1,\ldots,k$; (105) is by considering the "time-sharing" random variable q taking values over $1,2,\ldots,n$ uniformly, such that $\mathbf{x}_{u,q}$, $\mathbf{x}_{u^c,q}$, and $\mathbf{y}_{u^c,q}$ are the corresponding uniform mixtures of set of variables $\{\mathbf{x}_{u,1},\ldots,\mathbf{x}_{u,n}\}$, $\{\mathbf{x}_{u^c,1},\ldots,\mathbf{x}_{u^c,n}\}$, and $\{\mathbf{y}_{u^c,1},\ldots,\mathbf{y}_{u^c,n}\}$, respectively; (106) is by noting that, conditioned on $\mathbf{x}_{u^c,q}$ and $H(q)$, we have the Markov chain $q-\mathbf{x}_{u,q}-\mathbf{y}_{u^c,q}$ since $\mathbf{y}_{u^c,t}$ is completely determined by $\mathbf{x}_{u,t}$ and noise [according to (2)] independent of time t; and, finally, (107) is by dropping the index q everywhere and noting that the new variables still satisfy the relationship given in (2). Observe that if the supremum of the mutual information in the right-hand side of (107) is zero, then the mutual information term in the right-hand side of (93) is also zero which means that no reliable communication (i.e., $\epsilon\to 0$) is possible for large message sets ($\log_2 M>1$).

Since inequality (107) is valid for all cuts $u\in U$, we can write

$$(1-\epsilon)\log_2 M\leq\min_{u\in U} n\sup_{\substack{P_{\mathbf{x}}:\\ \mathbb{E}[|x_i|^2]\leq E_i/n}} I(\mathbf{x}_u;\mathbf{y}_{u^c}|\mathbf{x}_{u^c},\mathbf{H})+1. \quad (108)$$

Therefore, the energy per bit of the code is

$$\frac{E_{\text{total}}}{\log_2 M} \geq \frac{E_{\text{total}}}{n\min_{u\in U}\sup_{\substack{P_{\mathbf{x}}:\\ \mathbb{E}[|x_i|^2]\leq E_i/n}} I(\mathbf{x}_u;\mathbf{y}_{u^c}|\mathbf{x}_{u^c},\mathbf{H})} \times\left(1-\epsilon-\frac{1}{\log_2 M}\right) \quad (109)$$

$$=\left(\frac{\min_{u\in U}\sup_{\substack{P_{\mathbf{x}}:\\ \mathbb{E}[|x_i|^2]\leq E_i/n}} I(\mathbf{x}_u;\mathbf{y}_{u^c}|\mathbf{x}_{u^c},\mathbf{H})}{\sum_{i=1}^{k}E_i/n}\right)^{-1} \times\left(1-\epsilon-\frac{1}{\log_2 M}\right) \quad (110)$$

$$\geq\left(\sup_{\substack{p_1,\ldots,p_k\geq 0:\\ \sum_{i=1}^{k}p_i>0}}\min_{u\in U}\frac{\sup_{\substack{P_{\mathbf{x}}:\\ \mathbb{E}[|x_i|^2]\leq p_i}} I(\mathbf{x}_u;\mathbf{y}_{u^c}|\mathbf{x}_{u^c},\mathbf{H})}{\sum_{i=1}^{k}p_i}\right)^{-1} \times\left(1-\epsilon-\frac{1}{\log_2 M}\right) \quad (111)$$

where (110) is obtained from (92); and, to get (111), we have substituted E_i/n by p_i in (110) and taken supremum over p_i for all $i=1,\ldots,k$.

Recall that if E_b is ϵ-achievable, then for all $\delta > 0$

$$\lim_{M\to\infty} \frac{E_{\text{total}}}{\log_2 M} < E_b + \delta. \tag{112}$$

This implies that

$$\lim_{M\to\infty} \sup \left(\sup_{p_1,\ldots,p_k} \min_{u} \frac{\sup_{\substack{P_{\mathbf{x}}:\\ \mathbb{E}[|x_i|^2]\le p_i}} I(\mathbf{x}_u; \mathbf{y}_{u^c} | \mathbf{x}_{u^c}, \mathbf{H})}{\sum_{i=1}^k p_i} \right)^{-1} \times \left(1 - \epsilon - \frac{1}{\log_2 M}\right) \le E_b. \tag{113}$$

Moreover, if E_b is an achievable energy per bit value, then we can take supremum of the left-hand side of (113) over all $0 < \epsilon < 1$ to get

$$E_b \ge \left(\sup_{\substack{p_1,\ldots,p_k \ge 0\\ \sum_{i=1}^k p_i > 0}} \min_{u\in U} \frac{\sup_{\substack{P_{\mathbf{x}}:\\ \mathbb{E}[|x_i|^2]\le p_i}} I(\mathbf{x}_u; \mathbf{y}_{u^c} | \mathbf{x}_{u^c}, \mathbf{H})}{\sum_{i=1}^k p_i} \right)^{-1} \tag{114}$$

$$\ge \inf_{\substack{p_1,\ldots,p_k \ge 0\\ \sum_{i=1}^k p_i > 0}} \max_{u\in U} \frac{\sum_{i=1}^k p_i}{\sup_{\substack{P_{\mathbf{x}}:\\ \mathbb{E}[|x_i|^2]\le p_i}} I(\mathbf{x}_u; \mathbf{y}_{u^c} | \mathbf{x}_{u^c}, \mathbf{H})} \tag{115}$$

where (115) is obtained from (114) by noting that $\inf_{x\in\mathcal{X}} (f(x))^{-1} = (\sup_{x\in\mathcal{X}} f(x))^{-1}$ for any function f taking positive real values (and similarly, $\max_{y\in\mathcal{Y}} (g(y))^{-1} = (\min_{y\in\mathcal{Y}} g(y))^{-1}$ for any function g taking positive real values).

Therefore, $E_{b\min}$ should also satisfy (115). □

Appendix II
Proof of Corollary 2

Proof: The proof is by modifying the linear program (8) so that the number of constraints is reduced. Then, the dual of the resulting program is considered and its value evaluated for a specific variable vector.

Construct a new matrix $\tilde{\mathbf{L}}$ by reducing the u rows of $\mathbf{L}$ to $|\mathcal{S}|$ rows. In particular, we choose those rows which correspond to the cut S_i^c for all $S_i \in \mathcal{S}$. Thus, the (i,j)th coefficient of $\tilde{\mathbf{L}}$ is $\sum_{\ell\in S_i} g_{j\ell}$ if $j \in S_i^c$ and is 0 if $j \notin S_i^c$.

Since all the entries of S_i are from $\mathcal{R}$, S_i^c is a valid cut, and therefore, the rows of $\tilde{\mathbf{L}}$ are a subset of the rows of $\mathbf{L}$. Therefore, the value of the following linear program:

$$\tilde{E}_1 \triangleq \min_{\mathbf{q}\in\mathbb{R}^k} \mathbf{1}_k^T \mathbf{q} : \quad \tilde{\mathbf{L}}\mathbf{q} \ge (N_0 \log_e 2) \cdot \mathbf{1}_{|\mathcal{S}|}, \quad \mathbf{q} \ge \mathbf{0}_k \tag{116}$$

satisfies $E_1 \ge \tilde{E}_1$, which implies, from Theorem 1

$$E_{b\min} \ge \tilde{E}_1. \tag{117}$$

From the duality of linear programs, the value of (116) matches that of its dual, i.e.,

$$\tilde{E}_1 = (N_0 \log_e 2) \max_{\mathbf{q}\in\mathbb{R}^{|\mathcal{S}|}} \mathbf{1}_{|\mathcal{S}|}^T \mathbf{q} : \quad \tilde{\mathbf{L}}^T \mathbf{q} \le \mathbf{1}_k, \quad \mathbf{q} \ge \mathbf{0}_{|\mathcal{S}|}. \tag{118}$$

Next, it is easily verified that the vector

$$\mathbf{q} = \frac{1}{\gamma} \mathbf{1}_{|\mathcal{S}|} \tag{119}$$

satisfies the constraints of (118) for the choice of

$$\gamma = \max_{i=1,\ldots,k} \sum_{S\in\mathcal{S}\setminus i} \sum_{\ell\in S} g_{i\ell}. \tag{120}$$

Since $\tilde{E}_1$ is the maximum over all variable vectors satisfying the constraints of (118), substituting $\mathbf{q}$ in (118) from (119), we get

$$\tilde{E}_1 \ge (N_0 \log_e 2) \frac{|\mathcal{S}|}{\gamma}. \tag{121}$$

From (117), the right-hand side of (121) is still a lower bound on $E_{b\min}$. The conclusion in Corollary 2 is now immediate by evaluating the right-hand side of (121) using the value of γ from (120). □

Appendix III
Minimum Energy Per Bit of a Three-Terminal Network

We begin by reducing the number of variables in (29) to get

$$G(g_{12}, g_{13}, g_{23}, g_{32}) = \max_{\substack{\alpha_1,\alpha_2\ge 0:\\ \alpha_1+\alpha_2\le 1}} \min\{I_1, I_2, I_3\} \tag{122}$$

where

$$I_1(\alpha_1, \alpha_2) \triangleq (g_{12} + g_{13})\alpha_1 \tag{123}$$
$$I_2(\alpha_1, \alpha_2) \triangleq g_{12}\alpha_1 + g_{32}(1 - \alpha_1 - \alpha_2) \tag{124}$$
$$I_3(\alpha_1, \alpha_2) \triangleq g_{13}\alpha_1 + g_{23}\alpha_2. \tag{125}$$

Note that

$$I_1 \lesseqgtr I_2 \Longrightarrow \alpha_1 \lesseqgtr \frac{g_{32}}{g_{13} + g_{32}}(1 - \alpha_2) \tag{126}$$
$$I_2 \lesseqgtr I_3 \Longrightarrow (g_{12} - g_{32} - g_{13})\alpha_1 + g_{32} \lesseqgtr (g_{23} + g_{32})\alpha_2 \tag{127}$$
$$I_3 \lesseqgtr I_1 \Longrightarrow \alpha_2 \lesseqgtr \frac{g_{12}}{g_{23}}\alpha_1. \tag{128}$$

Next, we evaluate $G = G(g_{12}, g_{13}, g_{23}, g_{32})$ for different cases of channel gains, in the following manner. For each case of channel gains, we first find out the set of values (α_1, α_2) for which each of the terms I_1, I_2, and I_3 are minimized. Let $(\alpha_1, \alpha_2)_1$ be the value among all the values of (α_1, α_2) for which $I_1 \le I_2$ and $I_1 \le I_3$, and for which I_1 is maximized. Similarly, $(\alpha_2, \alpha_2)_2$ and $(\alpha_2, \alpha_2)_3$ corresponding to I_2 and

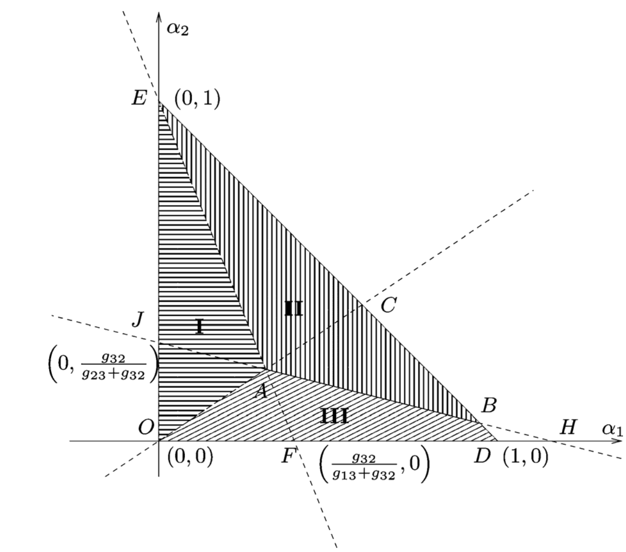

Fig. 8. Solution possibilities for (α_1, α_2) when $g_{13} + g_{32} \geq g_{12}$.

I_3, respectively, are determined. Finally, for the given case of channel gains g_{12}, g_{13}, g_{23}, and g_{32}, we have

$$G = \max\{I_1((\alpha_1, \alpha_2)_1), I_2((\alpha_1, \alpha_2)_2), I_3((\alpha_1, \alpha_2)_3)\}. \tag{129}$$

Now, let us calculate G for the different cases of channel gains as follows.

1) $g_{12} \geq g_{13}$ and $g_{13} + g_{32} \geq g_{12}$

Note that (α_1, α_2) can only take values in the triangle EOD as shown in Fig. 8. Marked also in the figure are the regions for which either I_1, I_2, or I_3 are the minimum of all three terms. In particular, from (126) and (128), for triangle EOA, $I_1 \leq I_2, I_3$. From (127) and (126), for triangle EAB, $I_2 \leq I_1, I_3$. From (127) and (128), for quadrilateral $OABD$, $I_3 \leq I_1, I_2$. The point

$$A \triangleq \left(\frac{g_{23}g_{32}}{g_{23}g_{32}+g_{13}g_{23}+g_{12}g_{32}}, \frac{g_{12}g_{32}}{g_{23}g_{32}+g_{13}g_{23}+g_{12}g_{32}}\right) \tag{130}$$

is common to all three regions.

A few remarks about points F, J, and H: Point $F \triangleq (g_{32}/(g_{32} + g_{13}), 0)$ always lies between O and D, where $O \triangleq (0, 0)$ and $D \triangleq (1, 0)$. Similarly, point $J \triangleq (0, g_{32}/(g_{32} + g_{23}))$ always lies between O and $E \triangleq (0, 1)$. Due to the defining conditions of this case, point $H \triangleq (g_{32}/(g_{13} + g_{32} - g_{12}), 0)$ always lies beyond D. Also, points B and C are given by $B \triangleq (g_{23}/(g_{12} + g_{23} - g_{13}), (g_{12} - g_{13})/(g_{12} + g_{23} - g_{13}))$ and $C \triangleq (g_{23}/(g_{12} + g_{23}), g_{12}/(g_{12} + g_{23}))$.

Next, note that in region EOA, maximizing $I_1 = (g_{12} + g_{13})\alpha_1$ implies maximizing α_1, which in turn implies operating at point A, i.e.,

$$I_1((\alpha_1, \alpha_2)_1) = \frac{(g_{12} + g_{13})g_{23}g_{32}}{g_{23}g_{32} + g_{13}g_{23} + g_{12}g_{32}}. \tag{131}$$

In the region $OABD$, we have to maximize

$$I_3((\alpha_1, \alpha_2)_3) = (g_{13}, g_{23}) \cdot (\alpha_1, \alpha_2)_3. \tag{132}$$

Since (g_{13}, g_{23}) lies in the first quadrant only, the maximum value of the inner product on the right-hand side of (132) is maximized at either A, B, or D.

Thus, for $g_{13} \geq g_{23}$, the optimal operating point is $(\alpha_1, \alpha_2)_3 = D$, which implies

$$I_3((\alpha_1, \alpha_2)_3) = g_{13}. \tag{133}$$

For $g_{13} \leq g_{23}$ and $g_{32}g_{23} \leq g_{12}g_{23} + g_{32}g_{13}$, $(\alpha_1, \alpha_2)_3 = B$, thus

$$I_3((\alpha_1, \alpha_2)_3) = \frac{g_{12}g_{23}}{g_{12} + g_{23} - g_{13}}. \tag{134}$$

For $g_{12}g_{23} + g_{32}g_{13} \leq g_{32}g_{23}$, $(\alpha_1, \alpha_2)_3 = A$, thus

$$I_3((\alpha_1, \alpha_2)_3) = \frac{(g_{12} + g_{13})g_{23}g_{32}}{g_{23}g_{32} + g_{13}g_{23} + g_{12}g_{32}}. \tag{135}$$

In region EAB, we have to maximize

$$I_2((\alpha_1, \alpha_2)_2) = (g_{12} - g_{32}, -g_{32}) \cdot (\alpha_1, \alpha_2)_2 + g_{32} \tag{136}$$

where $(\alpha_1, \alpha_2)_2$ can only be one of points E, A, and B depending upon the channel gains. Note that vector $(g_{12} - g_{32}, -g_{32})$ lies only in third and fourth quadrants.

For $g_{32}g_{23} \leq g_{12}g_{23} + g_{32}g_{13}$, $(\alpha_1, \alpha_2)_2 = B$, thus

$$I_2((\alpha_1, \alpha_2)_2) = \frac{g_{12}g_{23}}{g_{12} + g_{23} - g_{13}}. \tag{137}$$

For $g_{12}g_{23} + g_{32}g_{13} \leq g_{32}g_{23}$, $(\alpha_1, \alpha_2)_3 = A$, thus

$$I_2((\alpha_1, \alpha_2)_2) = \frac{(g_{12} + g_{13})g_{23}g_{32}}{g_{23}g_{32} + g_{13}g_{23} + g_{12}g_{32}}. \tag{138}$$

The conditions for operating at point E are never satisfied. Before moving further, we note the following implications:

$$g_{13} \lesseqgtr \frac{g_{12}g_{23}}{g_{12} + g_{23} - g_{13}} \\ \Downarrow \\ g_{13} \lesseqgtr g_{23} \tag{139}$$

Also

$$\frac{(g_{12} + g_{13})g_{23}g_{32}}{g_{23}g_{32} + g_{12}g_{32} + g_{13}g_{23}} \lesseqgtr \frac{g_{12}g_{23}}{g_{12} + g_{23} - g_{13}} \\ \Downarrow \\ g_{32}g_{23} \lesseqgtr g_{12}g_{23} + g_{13}g_{32}. \tag{140}$$

Finally, for each case of channel gain, we need to maximize G over the maximal values within the three regions. Putting together our analysis, the following subcases of channel gains emerge.

a) $g_{23} \leq g_{13}$. For this case

$$G = \max\left\{\frac{(g_{12} + g_{13})g_{23}g_{32}}{g_{23}g_{32} + g_{13}g_{23} + g_{12}g_{32}}, g_{13}, \frac{g_{12}g_{23}}{g_{12} + g_{23} - g_{13}}\right\}. \tag{141}$$

From (139) and (140), (141) implies

$$G = g_{13}. \tag{142}$$

b) $g_{13} \le g_{23} \le \frac{g_{12}}{g_{32}} g_{23} + g_{13}$. For this case

$$G = \max\left\{ \frac{(g_{12}+g_{13})g_{23}g_{32}}{g_{23}g_{32}+g_{13}g_{23}+g_{12}g_{32}}, \frac{g_{12}g_{23}}{g_{12}+g_{23}-g_{13}} \right\}. \tag{143}$$

From (140), (143) implies

$$G = \frac{g_{12}g_{23}}{g_{12}+g_{23}-g_{13}}. \tag{144}$$

c) $\frac{g_{12}}{g_{32}} g_{23} + g_{13} \le g_{23}$. For this case, maximal value within all regions is the same. Therefore

$$G = \frac{(g_{12}+g_{13})g_{23}g_{32}}{g_{23}g_{32}+g_{13}g_{23}+g_{12}g_{32}}. \tag{145}$$

2) $g_{12} \ge g_{13}$ and $g_{13} + g_{32} \le g_{12}$. The analysis of case 2) is almost the same as that of case 1). The regions where I_1, I_2, and I_3 are minimum are the same as before. The only difference from the solution region of the previous case is that the point H now lies to the left of the origin. For this case, in region EOA, where term I_1 is the least, operating at point A maximizes I_1. Therefore, $I_1((\alpha_1, \alpha_2)_1)$ is as given in (131). In region $OABD$, if $g_{13} \ge g_{23}$, the optimal operating point is at D, giving (133); if $g_{13} \le g_{23}$, the optimal operating point is B giving (134). Note that since $g_{12} \ge g_{32}$, condition $g_{32}g_{23} \le g_{12}g_{23} + g_{32}g_{13}$ always holds, so we never operate at point A. As before, in region EAB, we only operate at point B giving us (137). We do not operate at point A since condition $g_{32}g_{23} \le g_{12}g_{23} + g_{32}g_{13}$ always holds.
Therefore, the following subcases of the channel gains emerge.

a) $g_{23} \le g_{13}$. For this case

$$G = \max\left\{ \frac{(g_{12}+g_{13})g_{23}g_{32}}{g_{23}g_{32}+g_{13}g_{23}+g_{12}g_{32}}, g_{13}, \frac{g_{12}g_{23}}{g_{12}+g_{23}-g_{13}} \right\}. \tag{146}$$

So, from (139) and (140), (146) implies

$$G = g_{13}. \tag{147}$$

b) $g_{13} \le g_{23}$. For this case

$$G = \max\left\{ \frac{(g_{12}+g_{13})g_{23}g_{32}}{g_{23}g_{32}+g_{13}g_{23}+g_{12}g_{32}}, \frac{g_{12}g_{23}}{g_{12}+g_{23}-g_{13}} \right\}. \tag{148}$$

So, from (140), (148) implies

$$G = \frac{g_{12}g_{23}}{g_{12}+g_{23}-g_{13}}. \tag{149}$$

Appendix IV
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

$$\tilde{q}_{j(t)} \le (k-1)\, q_{j(t)} \tag{150}$$

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

Now we go to 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. So, $S(2) = \{1\}$. For any $2 \le t \le k$, suppose that the nodes in $S(t)$ have all decoded the message. We claim that some node [node $j(t)$] has gathered enough energy ($\ge N_0 \log_e 2$) by the end of slot $t-1$ for it to decode the message reliably. Suppose otherwise, i.e., for all $j \in S(t)^c \triangleq \{1, \dots, k\} \setminus S(t)$, we have

$$\sum_{\ell \in S(t)} \tilde{q}_\ell\, g_{\ell j} < N_0 \log_e 2. \tag{151}$$

This implies

$$\sum_{j \in S(t)^c} \sum_{\ell \in S(t)} \tilde{q}_\ell\, g_{\ell j} < (k-t+1)\, N_0 \log_e 2. \tag{152}$$

However, by our update rule for $\tilde{q}_{j(t)}$, for all $\ell \in S(t)$, $\tilde{q}_\ell \ge (k-t+1)\, q_\ell$. Therefore, from (152)

$$\sum_{j \in S(t)^c} \sum_{\ell \in S(t)} q_\ell\, g_{\ell j} < N_0 \log_e 2. \tag{153}$$

Recall that $\mathbf{q}$ satisfies the constraints in (8). Since $S(t)$ is a valid cut, consider the row in $\mathbf{L}$ corresponding to the cut $S(t)$. The constraint due to $S(t)$ dictates that

$$\sum_{\ell \in S(t)} q_\ell \sum_{j \in S(t)^c} g_{\ell j} \ge N_0 \log_e 2 \tag{154}$$

which directly contradicts (153). Therefore, there exists 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 holds for all $t = 2, \dots, k$. □

Appendix V
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 (8) dictates that

$$g_{12}\, q_1 \ge N_0 \log_e 2 \tag{155}$$

since g_{i2} is zero for $i = 2, 3, \ldots, k$. 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 $2 \le t \le k-1$, suppose that nodes $2, \ldots, t$ decode the message successfully by the end of time slot $t-1$. Consider the cut which isolates node $t+1$, i.e., the cut $\{1, 2, \ldots, t, t+2, \ldots, k\}$. This is a valid cut which appears as some row in $\mathbf{L}$ in (8). Therefore

$$\sum_{\substack{i=1\\ i\neq t+1}}^{k} q_i\, g_{i(t+1)} \ge N_0 \log_e 2. \tag{156}$$

Since we have the topological sort on the set of nodes, $g_{i(t+1)} = 0$ for $i > t+1$. Thus, (156) is equivalent to

$$\sum_{i=1}^{t} q_i\, g_{i(t+1)} \ge N_0 \log_e 2. \tag{157}$$

Since the node set $1, 2, \ldots, t$ has 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. □

Appendix VI
Proof of Lemma 4

Proof: Part 1 is directly by construction in Algorithm 2, since if a node in cluster S belongs to a subcell a, then there is no node in the eight subcells adjacent to a which are not a part of the cluster S. Thus, the minimum separation between a node belonging to S and any node not belonging to S is the side length $s/(9\bar{k}+1)$ of a subcell.

Proof of part 2 is by contradiction. Suppose that the cluster S_{m_1,m_2} spills over an adjacent supercell, i.e., without loss of generality suppose that the cluster $S_{0,0}$ contains a node belonging to cell $C(3,1)$ which belongs to a different supercell [one whose leader cell is $C(4,1)$]. By our construction of clusters, this implies that there is a path of adjacent (vertically, horizontally, or diagonally) nonempty subcells that spans from the boundary of $C(1,1)$ to the boundary of $C(3,1)$. Such a path would consist of at least $9\bar{k}+1$ subcells. Since the maximum number of nodes in the supercell with the leader cell $C(1,1)$ is $9\bar{k}$, this contradicts the statement that each subcell in the path is nonempty. This proves part 2.

Since cluster S_{m_1,m_2} is confined to a single supercell, and each supercell forms exactly one cluster, the number of clusters is exactly $n^2/9$, thereby proving part 3. Furthermore, the number of nodes in S_{m_1,m_2} is at least the number of nodes in the cell $C(3m_1+1, 3m_2+1)$ and at most the number of nodes in the corresponding supercell. This proves part 4. □

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] E. Belmega, B. Djeumou, and S. Lasaulce, "Gaussian broadcast channels with an orthogonal and bidirectional cooperation link," *EURASIP J. Wireless Commun. Netw.*, vol. 2008, no. 4, 2008, DOI: 10.1155/2008/341726.

[3] M. Čagalj, J. Hubaux, and C. Enz, "Minimum-energy broadcast in all-wireless networks: NP-completeness and distribution issues," in *Proc. 8th Annu. Int. Conf. Mobile Comput. Netw.*, Atlanta, GA, Sep. 2002, pp. 172–182.

[4] G. Caire, D. Tuninetti, and S. Verdú, "Suboptimality of TDMA in the low-power regime," *IEEE Trans. Inf. Theory*, vol. 50, no. 4, pp. 608–620, Apr. 2004.

[5] T. Cormen, C. Leiserson, R. Rivest, and C. Stein, *Introduction to Algorithms*, 2nd ed. Cambridge, MA: MIT Press, 2001.

[6] 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.

[7] T. Cover and J. Thomas, *Elements of Information Theory*, 2nd ed. New York: Wiley, 2006.

[8] R. Dabora and S. D. Servetto, "Broadcast channels with cooperating decoders," *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5438–5454, Dec. 2006.

[9] A. E. Gamal, M. Mohseni, and S. Zahedi, "Bounds on capacity and minimum energy-per-bit for AWGN relay channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1545–1561, Apr. 2006.

[10] 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.

[11] A. Jain, S. Kulkarni, and S. Verdú, "Multicasting in large wireless networks: Bounds on the minimum energy per bit," *IEEE Trans. Inf. Theory*, vol. 57, no. 1, pp. 14–32, Jan. 2011.

[12] 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 Annu. Conf. Commun. Control Comput.*, Monticello, IL, 2009, pp. 740–747.

[13] D. E. Knuth, *The Art of Computer Programming*. Reading, MA: Addison-Wesley, 1969, vol. 1.

[14] 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.

[15] A. Lapidoth, I. Telatar, and R. Urbanke, "On wide-band broadcast channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 12, pp. 3250–3258, Dec. 2003.

[16] S. Lasaulce and A. Klein, "Gaussian broadcast channels with cooperating receivers: The single common message case," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, Toulouse, France, May 2006, vol. 4, DOI: 10.1109/ICASSP.2006.1660901.

[17] W. Liang, "Constructing minimum-energy broadcast trees in wireless ad hoc networks," in *Proc. 3rd ACM Int. Symp. Mobile Ad Hoc Netw. Comput.*, New York, 2002, pp. 112–122.

[18] Y. Liang and V. V. Veeravalli, "Cooperative relay broadcast channels," *IEEE Trans. Inf. Theory*, vol. 53, no. 3, pp. 900–928, Mar. 2007.

[19] 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.

[20] 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.

[21] C. T. K. Ng, I. Maric, A. J. Goldsmith, S. Shamai, and R. D. Yates, "Iterative and one-shot conferencing in relay channels," in *Proc. IEEE Inf. Theory Workshop*, Punta del Este, Uruguay, Mar. 2006, pp. 193–197.

[22] A. S. Tanenbaum, *Computer Networks*, 4th ed. Upper Saddle River, NJ: Prentice-Hall, 2002.

[23] S. Verdú, "Spectral efficiency in the wideband regime," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1319–1343, Jun. 2002.

[24] S. Verdú, "On channel capacity per unit cost," *IEEE Trans. Inf. Theory*, vol. 36, no. 5, pp. 1019–1030, Sep. 1990.

[25] 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 Annu. Joint Conf. IEEE Comput. Commun. Soc.*, Tel-Aviv, Israel, 2000, vol. 2, pp. 585–594.

[26] Y. Yao, X. Cai, and G. B. Giannakis, "On energy efficiency and optimum resource allocation of relay transmissions in the low-power regime," *IEEE Trans. Wireless Commun.*, vol. 4, no. 6, pp. 2917–2927, Nov. 2005.

Aman Jain received the B.Tech. degree from Indian Institute of Technology—Kanpur, India, in 2005 and the M.A. and Ph.D. degrees from Princeton University, Princeton, NJ, in 2007 and 2010, respectively, all in electrical engineering.

In 2009, he was a Graduate Fellow with the McGraw Center for Teaching and Learning at Princeton University. Since 2010, he has been with Goldman,

Sachs & Co. His research interests include information theory, wireless communications, networks, and signal processing.

Dr. Jain received the Outstanding Teaching Assistant Award from the Department of Electrical Engineering, Princeton University. In 2000, he attended the International Mathematics Olympiad Training Camp, India.

Sanjeev R. Kulkarni (M'91–SM'96–F'04) received the B.S. degree in mathematics, the B.S. degree in electrical engineering, and the M.S. degree in mathematics from Clarkson University, Potsdam, NY, in 1983, 1984, and 1985, respectively, the M.S. degree in electrical engineering from Stanford University, Stanford, CA, in 1985, and the Ph.D. degree in electrical engineering from the Massachusetts Institute of Technology (MIT), Cambridge, in 1991.

From 1985 to 1991, he was a Member of the Technical Staff at MIT Lincoln Laboratory. Since 1991, he has been with Princeton University, Princeton, NJ, where he is currently Professor of Electrical Engineering, and an affiliated faculty member in the Department of Operations Research and Financial Engineering and the Department of Philosophy. He spent January 1996 as a Research Fellow at the Australian National University, 1998 with Susquehanna International Group, and summer 2001 with Flarion Technologies. His research interests include statistical pattern recognition, nonparametric statistics, learning and adaptive systems, information theory, wireless networks, and image/video processing.

Prof. Kulkarni received an Army Research Office (ARO) Young Investigator Award in 1992 and a National Science Foundation (NSF) Young Investigator Award in 1994. He has served as an Associate Editor for the IEEE TRANSACTIONS ON INFORMATION THEORY.

Sergio Verdú (S'80–M'84–SM'88–F'93) received the Telecommunications Engineering degree from the Universitat Politècnica de Barcelona, Barcelona, Spain, in 1980 and the Ph.D. degree in electrical engineering from the University of Illinois at Urbana-Champaign, Urbana, in 1984.

Since 1984, he has been a member of the faculty of Princeton University, Princeton, NJ, where he is the Eugene Higgins Professor of Electrical Engineering.

Prof. Verdú is the recipient of the 2007 Claude E. Shannon Award and the 2008 IEEE Richard W. Hamming Medal. He is a member of the National Academy of Engineering and was awarded a Doctorate Honoris Causa from the Universitat Politècnica de Catalunya in 2005. He is a recipient of several paper awards from the IEEE: the 1992 Donald Fink Paper Award, the 1998 Information Theory Outstanding Paper Award, an Information Theory Golden Jubilee Paper Award, the 2002 Leonard Abraham Prize Award, the 2006 Joint Communications/Information Theory Paper Award, and the 2009 Stephen O. Rice Prize from IEEE Communications Society. He has also received paper awards from the Japanese Telecommunications Advancement Foundation and from Eurasip. In 1998, Cambridge University Press published his book *Multiuser Detection*, for which he received the 2000 Frederick E. Terman Award from the American Society for Engineering Education. He served as President of the IEEE Information Theory Society in 1997. He is currently Editor-in-Chief of *Foundations and Trends in Communications and Information Theory*.