Multicasting in Large Wireless Networks: Bounds on the Minimum Energy Per Bit

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

***Abstract*—In this paper, we consider scaling laws for maximal energy efficiency of communicating a message to all the nodes in a wireless network, as the number of nodes in the network becomes large. Two cases of large wireless networks are studied—dense random networks and constant density (extended) random networks. In addition, we also study finite size regular networks in order to understand how regularity in node placement affects energy consumption. We first establish an information-theoretic lower bound on the minimum energy per bit for multicasting in arbitrary wireless networks when the channel state information is not available at the transmitters. Upper bounds are obtained by constructing a simple flooding scheme that requires no information at the receivers about the channel states or the locations and identities of the nodes. The gap between the upper and lower bounds is only a constant factor for dense random networks and regular networks, and differs by a poly-logarithmic factor for extended random networks. Furthermore, we show that the proposed upper and lower bounds for random networks hold almost surely in the node locations as the number of nodes approaches infinity.**

***Index Terms*—Cooperative communication, minimum energy per bit, multicasting, wideband communication, wireless networks.**

I. Introduction

A. Prior Work

DETERMINING the energy efficiency of a point-to-point channel is a fundamental information-theoretic problem. While the minimum energy per bit requirement for reliable communication is known for a general class of channels [25], [26], the problem is considerably more complicated for networks. Even when just one helper (*relay*) node is added to the two terminal additive white Gaussian noise (AWGN) channel, the minimum energy per bit is still unknown, though progress has been made in [3], [9], and [28] among others. The minimum energy per bit for the general Gaussian multiple-access channel, the broadcast channel, and the interference channel has been considered in [4], [13], [25], and [26]. As the number of relays k in a network grows, it is interesting to analyze the improvement in energy efficiency as a function of k. It is shown in [7] that a two-hop *distributed beamforming* scheme is energy efficient for dense random networks, with the energy requirement falling as $\Theta(1/\sqrt{k})$. In this scheme, however, the relay nodes require knowledge of the channel states of the forward and backward links. The energy efficiency of multihopping in a unicast random network setting is considered in [8] and references therein. The impact of noncoherence and multicasting remains largely unexplored.

Cooperation between nodes (also known as *cooperative diversity*) leads to capacity or reliability gains even with simple communication schemes (e.g., see [12], [23], and [24] among others). One of the simple ideas for cooperation in a multicast setting is to let several nodes transmit the same signal (at lower power levels), so that each receiving node can combine several low reliability signals to construct progressively better estimates. This scheme works only in multicasting where all nodes retransmit the same message. Multistage decode-and-forward schemes to reduce the transmission energy in a network are proposed in [11], [15], and [22]. The optimization of a scheme of this nature can be formulated as an optimal cooperative broadcast [11], [19] or an *accumulative broadcast* problem [15]. In these formulations, an optimal transmission order for the nodes is constructed, and then given such an ordering, the transmission energy is minimized by solving a linear program for the power distribution, subject to the condition that each node receives a minimum amount of power.

The problem of multicasting with minimum energy consumption has drawn a lot of research interest. For the case of wired networks, the problem can be formulated as the well-known minimum cost spanning tree problem. However, for wireless networks, there is an inherent wireless multicast advantage [27] that allows all the nodes within the coverage range to receive the message at no additional cost. The minimum energy broadcast problem was formulated as a broadcast tree problem in [27]. The formulation based on wireless multicast advantage, however, still misses the advantage of overhearing other transmissions over the network. This advantage, which is important in multicasting, has been referred to as cooperative wireless advantage in [11]. This suggests that a more fundamental approach to the modeling and analysis of wireless networks may yield better results based on exploiting the broadcast nature of wireless communications.

The power efficiency of a cooperative decode-and-forward multicasting scheme for dense random networks has also been studied in noninformation-theoretic setting [19], where efficient achievability schemes are presented for dense networks only.

Manuscript received March 10, 2009; revised February 24, 2010; accepted May 25, 2010. Date of current version December 27, 2010. This work was supported in part by the National Science Foundation under Grants CCF-06-35154 and CNS-09-05398, the National Science Foundation Science & Technology Center under Grant CCF-0939370 and the U.S. Army Research Office under grant W911NF-07-1-0185. The material in this paper was presented in part at the IEEE International Symposium on Information Theory, Seoul, South Korea, June 2009.

The authors are with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544 USA (e-mail: amanjain@princeton.edu; kulkarni@princeton.edu; verdu@princeton.edu).

Communicated by M. Gastpar, Associate Editor for Shannon Theory.
Digital Object Identifier 10.1109/TIT.2010.2090228

In contrast, our interest is in the order of growth of energy requirements for both dense and extended networks. In addition to upper bounds on energy consumption, we also provide lower bounds thereby allowing comparison with the fundamental limits. Another major difference of this work from the previous works is our emphasis on minimal network information, which prohibits the use of centrally optimized policies.

Other related works include study of multicast capacity scaling laws for the dense and extended networks in [18] and [17], respectively. The question of multicast capacity for multihopping is addressed in [10] and [29]. Energy-efficient area coverage using a multistage decode-and-forward scheme is studied in [20] among others.

B. Summary of Results

In this work, our aim is to determine the maximum possible energy efficiency (i.e., minimum transmission energy per information bit) for multicasting in various wireless networks of practical interest, when there is no constraint on the bandwidth or delay. We put special emphasis on the particular multicasting setting (*broadcasting*) where all the nodes are interested in a common message. Besides developing suitable converse bounds, we also show how cooperative communication is instrumental to approach them.

Severely energy-constrained large wireless networks provide the main motivation for our work. We succinctly identify the main physical factors determining the energy efficiency in such networks. In this context, we are able to provide fundamental information-theoretic limits on maximum possible efficiency, which depend only on the physical description of the communication channels and not on the bandwidth or cooperation strategies. On the other hand, to achieve high energy efficiency, we restrict ourselves to only those schemes which do not have access to network information like node placements or channel states, etc., since such information is difficult and expensive to gather in real-world networks.

In Theorem 1, we present a lower bound on the energy requirement per information bit for multicasting in arbitrary wireless networks. This lower bound is shown to be proportional to the *effective loss* of the network, which is a fundamental property of the network and is completely determined by the channel attenuation between the nodes and the set of destination nodes. This bound is applicable whenever channel state information is not available at the transmitters.

For the achievability part, we analyze a simple technique—*flooding* [14], [21] based on repetition coding. The basic idea is to collect energy from multiple transmissions to reconstruct the original message [15], [19]. In our flooding algorithm, we assume no knowledge of the node locations, identities, or the channel states at the nodes. We show that just the information about the number of nodes, the area of the network, and the fading statistics is sufficient to achieve the same order of energy scaling as that of the schemes with considerably more knowledge. The converse and achievability bounds on the minimum energy requirement per information bit are then evaluated for two cases of large random networks and for finite regular networks.

The physical channel is modeled as a fading channel subject to Gaussian noise. We operate in the wideband regime, which is essential to maximize the energy efficiency in a point-to-point Gaussian channel [25]. Since there is no bandwidth constraint, we allot a separate wideband transmission channel to each node. The power-constrained, wideband multicast setting considered here is particularly relevant to sensor networks [1].

The gain between a pair of nodes is determined by the distance between the nodes according to a path loss model. Specifically, we model the strength of the fading coefficient between a pair of nodes at distance r with $\alpha > 2$. Furthermore, we assume that the gain never exceeds $\bar{g}$ for any distance between the nodes.

The different kinds of networks that we study are as follows.

- Large dense random networks, where the $k-1$ nonsource nodes are placed independently and uniformly over a square area of size A_k that increases as $o\left(k/\log k\right)$. The minimum energy per bit in this case scales linearly with A_k.
- Large extended random networks, where the $k-1$ nonsource nodes are placed independently and uniformly over a square area of size A_k that increases linearly with k. The minimum energy per bit is shown to be lower bounded by $\Omega(k)$, with the constant depending on the node density. The flooding algorithm achieves energy consumption within a poly-logarithmic factor of the lower bound.
- Finite regular networks, where the network is partitioned into square cells, each of which contains exactly one node. Furthermore, the nodes are confined to a certain fraction of area within these cells. In general, the lower bound on the minimum energy per bit can depend on both the number of nodes as well as the node density. However, the energy consumption of the flooding algorithm is always within a constant factor of the lower bound.

For the case of large networks, we are interested in the asymptotic analysis (as $k \to \infty$) of the upper and lower bounds on the energy requirement per bit. On the other hand, regular networks are studied for all values of $k \geq 2$.

The remainder of this paper is structured as follows. In Section II, we introduce the system model. In Section III, we prove a general result about the minimum energy requirement of multicasting in a wireless network. A flooding algorithm is introduced in Section IV. In Section V, we specify the path loss model and study dense and extended random networks. Finally, in Section VI, we study finite-sized regular networks.

II. System Model

A. Channel Model

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

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

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

B. Problem Setup

All the nodes in the network are identical and are assumed to have receiving, processing, and transmitting capabilities. The nodes can also act as relays to help out with the task of communicating a message to the whole network. The total energy consumption of the network is simply the sum of transmission energies at all the nodes. To define a multicast relay network, we extend the three terminal relay channel setting of [5] to include multiple relays and multiple destination nodes. An error is said to have occurred when any of the intended nodes fails to decode the correct message transmitted by the source.

Consider a code for the network, with block length n. For $i = 1, \ldots, k$, the codeword at node i is 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 (node 1) is $\mathcal{M} = \{1, 2, \ldots, M\}$, then the codeword $x_1^{(n)}(m)$ is determined by the message m chosen equiprobably from the message set. At any other node $i \in \{2, \ldots, k\}$, the codeword $x_i^{(n)}$ is a function of the channel outputs $y_i^{(n)} = (y_{i,1}, y_{i,2}, \ldots, y_{i,n})$ at the node. Due to causality, the tth symbol $x_{i,t}$ of $x_i^{(n)}$ is a function of the first $t-1$ inputs at the node, i.e., $x_{i,t} = x_{i,t}(y_i^{(t-1)})$. This function, which defines the input–output relation at a nonsource node, is also called the *relay function*.

At any nonsource node i, in addition to a relay function, there may also be a *decoding function* (depending on whether the node is a *destination node*) which decodes a message $\hat{m}_i \in \mathcal{M}$ based on the n channel outputs $y_i^{(n)}$ at the node. Therefore, $\hat{m}_i = \hat{m}_i(y_i^{(n)})$.

Suppose that only a subset $\mathcal{R} \subseteq \{2, \ldots, k\}$ (also called the *destination set*) of the nodes is interested in receiving the message from the source node. When $\mathcal{R}$ contains two or more nodes, it is called a *multicast* setting.

The probability of error of the code is defined as

$$P_e \triangleq \frac{1}{M} \sum_{m \in \mathcal{M}} P_e[m] \tag{2}$$

where

$$P_e[m] \triangleq P[\exists i \in \mathcal{R} : \hat{m}_i \neq m | m \text{ is the message}]. \tag{3}$$

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 big 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.,

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

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

$$E_b \triangleq \frac{E_{\text{total}}}{\log_2 M}. \tag{5}$$

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 $0 \leq \epsilon < 1$.

In [26], channel capacity per unit cost was defined for a channel without restrictions on the number of channel uses. Here, we are interested in the reciprocal of this quantity.

Definition: Given $0 \leq \epsilon < 1$, $E \in \mathbb{R}_+$ is an ϵ-achievable energy per bit if for every $\delta > 0$ and all sufficiently large M, an $(n, M, (E+\delta)\log_2 M, \epsilon)$ code exists.

E is an achievable energy per bit if it is ϵ-achievable energy per bit for all $0 < \epsilon < 1$. The minimum energy per bit $E_{b\min}$ is the infimum of all the achievable energy per bit values. Sometimes, we deal with the normalized (w.r.t. noise spectral density N_0) version of $E_{b\min}$, which is represented by $\frac{E_b}{N_0}_{\min}$.

Minimal Information Framework: We derive scaling results for minimum energy per bit for different classes of networks. For a given number of nodes k, each class of networks has a set of possible network realizations. Our aim is to achieve low energy consumption per bit using no information at the nodes about the actual network realization (i.e., node locations). In addition, we also assume that the nodes have no information about the channel states. All the nonsource nodes have the same relay and decoding functions.

Providing local or global information to the nodes enlarges the set of possible coding schemes. Our converse results allow coding schemes to rely on any such information except for channel state information at the transmitters.

III. A Lower Bound on the Minimum Energy Per Bit

In this section, in Theorem 1, we show an information-theoretic lower bound on the minimum energy per bit for multicast in an arbitrary network. The bound depends on the destination nodes and the channel gains, through effective network loss defined below. It holds for any communication scheme where channel states are not known at the transmitters.

Theorem 1: In a network with k nodes, where node 1 is the source node and the destination set is $\mathcal{R} \subset \{2, \ldots, k\}$, the required minimum transmit energy per bit satisfies

$$\frac{E_b}{N_0}_{\min}\left(\mathcal{R}\right) \geq \left(\frac{1}{G(\mathcal{R})}\right) \log_e 2 \tag{6}$$

where $1/G(\mathcal{R})$ (in decibels) is the effective network loss defined by the following relation:

$$G(\mathcal{R}) \triangleq \frac{1}{|\mathcal{R}|}\left(\max_{i\in\{1,\ldots,k\}} \sum_{j\in\mathcal{R}\setminus\{i\}} \mathbb{E}[|h_{ij}|^2]\right). \tag{7}$$

Before proving Theorem 1, we state Lemma 1 which provides a converse relating the minimum energy per bit to the channel capacity (see also [9] and [26]). Dropping the time indices, the channel (1) for the received symbol y_j at node j can be rewritten as

$$y_j = \mathbf{h}_j^T\mathbf{x} + z_j \tag{8}$$

where $\mathbf{x} = (x_1, \ldots, x_k)^T$ is the transmission symbol vector and $\mathbf{h}_j = (h_{1j}, \ldots, h_{(j-1)j}, 0, h_{(j+1)j}, \ldots, h_{kj})^T$ is the vector representing the fading. The complex Gaussian noise z_j is taken to be distributed according to $\mathcal{CN}(0, N_0)$.

Lemma 1: For the destination set $\mathcal{R}$, the minimum energy per bit for the network satisfies

$$E_{b\min}(\mathcal{R}) \geq \inf_{\substack{P_1,P_2,\ldots,P_k\geq 0\\ :\sum_{i=1}^k P_i>0}} \max_{j\in\mathcal{R}} \frac{\sum_{i=1}^k P_i}{\sup_{\substack{P_{\mathbf{x}}:\\ \mathbb{E}[|x_i|^2]\leq P_i \text{ for } i=1,\ldots,k}} I(\mathbf{x}; y_j|\mathbf{h}_j)}. \tag{9}$$

Proof: See Appendix I. □

A brief rationale for Lemma 1 is as follows. Pick a node j belonging to the destination set $\mathcal{R}$. For the given power constraints: $P_1, P_2, \ldots, P_k$ on the transmission power, consider the channel from the set of nodes $\{1, \ldots, j-1, j+1, \ldots, k\}$ to node j. By the *max-flow min-cut* bound, the rate of reliable communication to node j by the remaining nodes cannot exceed

$$C_j(P_1, \ldots, P_k) \triangleq \sup_{\substack{P_{\mathbf{x}}:\\ \mathbb{E}[|x_i|^2]\leq P_i \text{ for } i=1,\ldots,k}} I(\mathbf{x}; y_j|\mathbf{h}_j) \tag{10}$$

bits per channel use. Therefore, the number of channel uses per bit is at least $1/C_j(P_1, P_2, \ldots, P_k)$, which implies that the total energy spent per bit in communicating to node j is at least $\sum_{i=1}^k P_i/C_j$. While this energy is spent communicating with node j, the transmission may benefit other nodes as well. In general, all the other nodes may be able to decode the message just by listening to the transmissions intended for node j. However, the minimum energy required to communicate to node j does not exceed the minimum energy required to communicate to all the nodes in $\mathcal{R}$. Therefore, we can lower bound the total energy spent communicating to all the nodes in $\mathcal{R}$ by the energy spent communicating to any one of the nodes in $\mathcal{R}$.

Proof of Theorem 1: We can lower bound the minimum energy per bit by

$$\begin{aligned}
E_{b\min}(\mathcal{R}) &\geq \inf_{\substack{\boldsymbol{w}:\\ w_i\geq 0,\\ \sum_{i=1}^k w_i=1}} \inf_{P>0} \max_{j\in\mathcal{R}} \frac{P}{\left(\sup_{\substack{P_{\mathbf{x}}:\\ \mathbb{E}[|x_i|^2]\leq w_i P}} I(\mathbf{x}; y_j|\mathbf{h}_j)\right)} && (11)\\
&\geq \inf_{\substack{\boldsymbol{w}:\\ w_i\geq 0,\\ \sum_{i=1}^k w_i=1}} \max_{j\in\mathcal{R}} \inf_{P>0} \frac{P}{\left(\sup_{\substack{P_{\mathbf{x}}:\\ \mathbb{E}[|x_i|^2]\leq w_i P}} I(\mathbf{x}; y_j|\mathbf{h}_j)\right)} && (12)\\
&\geq \min_{\substack{\boldsymbol{w}:\\ w_i\geq 0,\\ \sum_{i=1}^k w_i=1}} \max_{j\in\mathcal{R}} \frac{N_0 \log_e 2}{\sum_{i=1}^k \mathbb{E}[|h_{ij}|^2] w_i} && (13)
\end{aligned}$$

where the explanation of the steps (11)–(13) is the following. The inequality (11) follows from Lemma 1 by rewriting it so that P is the total power and $\boldsymbol{w}$ is the fractional split of power among all the nodes. The bound in (12) follows from the fact that min–max is greater than or equal to max–min. To justify (13), note that the mutual information term in (12) corresponds to the capacity of a multiple transmit and single receive antenna system, which has been widely studied for additive Gaussian noise channels. We are interested in the case where channel state information is not available at the transmitters. For a given probability distribution on $\mathbf{x}$ (independent of $\mathbf{h}_j$), we can bound the mutual information in (12) as

$$\begin{aligned}
I(\mathbf{x}; \mathbf{h}_j^T\mathbf{x}+z_j|\mathbf{h}_j) &\leq \mathbb{E}\left[\log_2\left(1+\frac{1}{N_0}\mathbb{E}[|\mathbf{h}_j^T\mathbf{x}|^2|\mathbf{h}_j]\right)\right] && (14)\\
&\leq \frac{\log_2 e}{N_0}\mathbb{E}[|\mathbf{h}_j^T\mathbf{x}|^2] && (15)\\
&\leq \frac{\log_2 e}{N_0}\sum_{i=1}^k \mathbb{E}[|h_{ij}|^2] w_i P. && (16)
\end{aligned}$$

Note that, given $\mathbf{h}_j$, a constraint on the output $y_j = \mathbf{h}_j^T\mathbf{x}+z_j$ is that $\mathbb{E}[|y_j|^2] \leq \mathbb{E}[|\mathbf{h}_j^T\mathbf{x}|^2|\mathbf{h}_j] + N_0$; thus, the mutual information in (12) is maximized when y_j is Gaussian distributed with the given power constraint, which leads to (14); the bound in (15) is obtained using the simple fact that $\log_e(1+x) \leq x$ for all $x > 0$; we obtain (16) by maximizing the right-hand side of (15) among all $\mathbf{x}$ independent of $\mathbf{h}_j$ such that $\mathbb{E}[|x_i|^2] \leq w_i P$, taking into account the fact that the channel coefficients are independent with zero mean. Note that (16) directly implies (13) through (12).

Now that we have established (13) note that

$$\begin{aligned}
&\max_{\substack{\boldsymbol{w}:\\ w_i\geq 0,\\ \sum_{i=1}^k w_i=1}} \min_{j\in\mathcal{R}} \sum_{i=1}^k \mathbb{E}[|h_{ij}|^2] w_i\\
&\quad\leq \max_{\substack{\boldsymbol{w}:\\ w_i\geq 0,\\ \sum_{i=1}^k w_i=1}} \frac{1}{|\mathcal{R}|}\sum_{j\in\mathcal{R}}\sum_{i=1}^k \mathbb{E}[|h_{ij}|^2] w_i && (17)\\
&\quad= \max_{\substack{\boldsymbol{w}:\\ w_i\geq 0,\\ \sum_{i=1}^k w_i=1}} \frac{1}{|\mathcal{R}|}\sum_{i=1}^k w_i\left(\sum_{j\in\mathcal{R}\setminus\{i\}} \mathbb{E}[|h_{ij}|^2]\right) && (18)\\
&\quad= \frac{1}{|\mathcal{R}|}\max_{i\in\{1,\ldots,k\}}\left(\sum_{j\in\mathcal{R}\setminus\{i\}} \mathbb{E}[|h_{ij}|^2]\right) && (19)\\
&\quad= G(\mathcal{R}) && (20)
\end{aligned}$$

$\texttt{FLOOD}(E_{b1}, E_{b2})$

1) The source node transmits only in the first 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 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. 1. The flooding algorithm: $\texttt{FLOOD}(E_{b1}, E_{b2})$.

where (17) is obtained by upper-bounding the minimum by the average; the maximum in (18) is attained when all the weight is put on that i for which $\sum_j \mathbb{E}[|h_{ij}|^2]$ is largest. Substituting $G(\mathcal{R})$ from (20) in (13) provides the requisite lower bound (6) on the minimum energy per bit. □

Remark 1: While we expect the minimum energy per bit to increase as the destination set becomes larger, it can be shown that the effective network loss does not always increase with the size of destination set. Therefore, it is useful to maximize the right-hand side of (6) by considering all nonempty subsets of the destination set $\mathcal{R}$. Thus, a tighter bound on the minimum energy per bit is given by

$$\frac{E_b}{N_0}_{\min}\left(\mathcal{R}\right) \geq \max_{\substack{\mathcal{R}' \subset \mathcal{R}, \\ \mathcal{R}' \neq \phi}} \frac{\log_e 2}{G(\mathcal{R}')}. \tag{21}$$

Remark 2: For the simple case of a point-to-point Gaussian channel, the effective network loss is simply the reciprocal of the channel gain from the source to the destination. The bound is tight in this case [25, Th. 1]. Theorem 1 can be thought of an extension of the point-to-point case since we assume in the derivation of (6) (in Lemma 1) that for any node $i \in \mathcal{R}$ the rest of the nodes cooperate perfectly to communicate information to node i. Thus, we account for the energy in the "last hop" to node i from the rest of the nodes.

While Theorem 1 holds for general multicasting, we are interested in broadcasting in the networks considered now onwards.

IV. Flooding Algorithm

To derive upper bounds on the minimum energy per bit for broadcasting, we use a version of the well-known flooding algorithm. The flooding in the network occurs on top of the wideband, decode-and-forward hops. These hops could be one-to-one, one-to-many, many-to-one, or even many-to-many. The flooding algorithm, with suitable parameter values, is used to achieve energy-efficient multicasting for the various networks considered later in the paper.

Since minimum energy per bit requires very small spectral efficiency even in the point-to-point case, we do not place any bandwidth constraints. Therefore, we can assign each transmitter its own wide frequency band. In the wideband regime, the knowledge of the channel states at the receiver does not decrease the minimum energy per bit [25]. Furthermore, a necessary condition for reliable decoding is that the received energy per bit be greater than $N_0 \log_e 2$. Various wideband communication schemes can be constructed which let the receivers reliably decode a message if the total received energy per bit exceeds $N_0 \log_e 2$ [25].

Note that in case there are multiple sources in the network, they can transmit their messages independently of other sources by separation in either time or frequency. In this case, the total energy consumption for all the sources is simply the sum of the energy consumption of broadcast by each source.

1) Description of the Algorithm: The flooding algorithm consists of two parts: an outer algorithm and an inner coding scheme. The outer algorithm $\texttt{FLOOD}(E_{b1}, E_{b2})$ is the description at the *time slot* level using the *decoding* and *encoding* functionalities provided by the inner scheme. See Fig. 1.

Time is divided into slots: $1, 2, \ldots, T$, each time slot consisting of enough time to let a node transmit one codeword. Multiple nodes can simultaneously transmit in a slot, albeit in their own mutually orthogonal frequency bands. The transmission process is initiated by the source node which is the only node transmitting in the first slot. In any slot thereafter, whether a nonsource node transmits and what it transmits is dependent on when and what the node has decoded so far. In particular, if a node decodes a message for the first time in slot t, it retransmits the codeword corresponding to the decoded message in slot $t+1$. The decoding process and the determination of the codeword to be transmitted is handled by the inner coding scheme.

Note that every node transmits either never or once. The total number of slots T in the algorithm is a design parameter that depends on the size of the network.

2) Energy Consumption of $\texttt{FLOOD}(E_{b1}, E_{b2})$: In a network with k nodes, the source node transmits energy per bit E_{b1} and each of the remaining $k-1$ nonsource nodes transmit either 0 or E_{b2}. Therefore, the total energy consumed per information bit by $\texttt{FLOOD}(E_{b1}, E_{b2})$ is at most

$$E_{b\text{total}} \leq E_{b1} + (k-1)E_{b2}. \tag{22}$$

Instead of a single flooding scheme, we will demonstrate a sequence of flooding schemes which achieve a vanishing probability of error. $E_{b\text{flood}}$ will be used to denote the infimum of $E_{b\text{total}}$ over this sequence of schemes. Clearly, $E_{b\text{flood}}$ is an achievable energy per bit for the network and thus an upper bound on the $E_{b\min}$ of the network.

3) Inner Coding Scheme: The *transmit* operation in $\texttt{FLOOD}(E_{b1}, E_{b2})$ uses identical codebooks for all nodes. The task for each decoder is to observe transmissions over multiple time slots and frequency bands. Using these observations, it forms a reliable estimate of the source message. At the end of each time slot it determines whether it has enough information to decode the message. If not, it keeps quiet and waits for more transmissions. If it is able to decode a message, it re-encodes the decoded message and transmits it in the next slot for the benefit of its peers, and remains quiet after that.

Remark 3: Note that FLOOD is a fair algorithm in the sense that all the nonsource nodes expend the same amount of energy. The uniform energy allocation is a natural policy since the nodes do not have much information about the network (such as, location of its neighbors or its own distance from the source) to regulate the transmission power. An advantage of uniform energy allocation is that the network lifetime is maximized since the energy consumption is evenly distributed over all the nonsource nodes. Also, we note in advance that in all the instances of FLOOD studied in this paper, while the source node may spend a different amount of energy, it is always greater than the energy expenditure at the nonsource nodes. This can be thought of as the cost of initiating a broadcast in the network.

V. Large Random Networks

This section is devoted to the analysis of random networks, where the number of nodes k goes to infinity and their locations are random. We focus on the cases of dense and extended networks. In each case, we obtain both upper and lower bounds on $E_{b\min}$ that hold almost surely in the network topology as $k \to \infty$.

In both cases, k nodes are placed over a square of area A_k. The diagonal coordinates of the square are $(0,0)$ and $(\sqrt{A_k}, \sqrt{A_k})$, and the source node is placed on the coordinate $(0,0)$. This is a least favorable location for the source node but turns out to be irrelevant for the scaling laws we derive.

A. Path Loss Model

The channel gain $\mathbb{E}[|h_{ij}|^2]$ of the link between nodes i and j is determined by their separation r_{ij}. This relation is given by a monotonically decreasing *power gain* or *path loss* function $g(r) : \mathbb{R}_+ \mapsto \mathbb{R}_+$, i.e.,

$$\mathbb{E}[|h_{ij}|^2] = g(r_{ij}) \tag{23}$$

where, for all $r \geq r_0$

$$g(r) = r^{-\alpha} \tag{24}$$

where $r_0 > 0$ and $\alpha > 2$ are constants of the model.

To deal with the near-field case, we also put an upper bound on the gain function, i.e., there is a constant $\bar{g} > 0$ such that

$$g(0) \leq \bar{g} \tag{25}$$

since the gain cannot be arbitrarily large. Thus, the path loss model is completely characterized by α, r_0, and $\bar{g}$.

B. Dense Random Networks

A dense random network with $k \geq 2$ nodes consists of a source node at the origin and $k-1$ nonsource nodes distributed independently and uniformly over a square of area

$$A_k = o\left(\frac{k}{\log k}\right). \tag{26}$$

In addition, we also assume that

$$r_0^2 \leq 8A_k \tag{27}$$

for all $k \geq 2$.

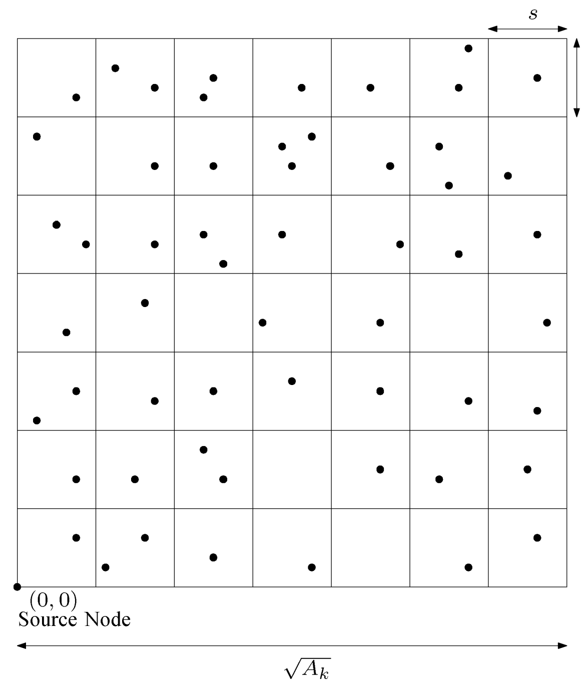

Fig. 2. Dense random network.

The results for this case are presented in Theorem 2 which states that the minimum energy per bit of a dense random network scales linearly with area, almost surely as $k \to \infty$, and does not depend directly on the actual number of nodes. The almost sure statement is made with respect to the location of the nodes.

Theorem 2: With probability 1, the node placement is such that

$$c_1 \leq \frac{1}{A_k} \frac{E_b}{N_0}_{\min} \tag{28}$$

and

$$\frac{1}{A_k} \frac{E_b}{N_0}_{\text{flood}} \leq c_2 \tag{29}$$

for all but a finite number of k, where

$$c_1 = \frac{2 \log_e 2}{49\, \bar{g} r_0^2 + \dfrac{2^{\alpha+2}}{\alpha-2} \dfrac{3}{r_0^{\alpha-2}}} \tag{30}$$

and

$$c_2 = 24\, r_0^{\alpha-2} \log_e 2. \tag{31}$$

Proof: We begin by partitioning the area A_k into square cells with side length

$$0 < s \leq \frac{r_0}{\sqrt{8}} \tag{32}$$

independent of k. See Fig. 2. Some of the cells may not be *whole*, i.e., they may not cover an area of s^2. However, all these cells would only lie along the upper and right sides of the square A_k. For simplicity, we restrict our attention to whole cells (i.e., $\sqrt{A_k}$ is a multiple of s) for the rest of the proof. Note that there are a total of A_k/s^2 cells in the network. We use "cell" to not only refer to the geographical cell but also to the set of nodes falling within the cell. Let $\mathcal{C}$ be the set of cells, and let $\nu(C)$ denote the number of nodes in cell $C \in \mathcal{C}$. We use *origin cell* to refer to the cell containing the source node, i.e., the cell with its lower left corner at $(0,0)$.

For any $\delta > 0$, define a *good placement* event $\mathcal{D}_k$ as the collection of node placement realizations for k nodes for which all the cells contain at least $(1-\delta)(k-1)s^2/A_k$ nodes and less than $((1+\delta)(k-1)s^2/A_k)+1$ nodes. Note that the good placement event is the one where each cell contains number of nodes approximately proportional to its area. Instrumental to both direct and converse parts of the proof, Lemma 2 lower bounds the probability of good placement for a given k.

Lemma 2:

$$P[\mathcal{D}_k^c] \leq \frac{2A_k}{s^2}\exp\left(-\delta^2(1-\delta)\frac{(k-1)s^2}{2A_k}\right) \tag{33}$$

for all $k \geq 2$.

Proof: See Appendix II. □

Let $\mathcal{F}_k$ be the collection of all node placement realizations for k nodes for which both (28) and (29) hold. To prove Theorem 2, we only need to prove that

$$P[\liminf_{k\to\infty} \mathcal{F}_k] = 1. \tag{34}$$

To do so, we show that any realization within $\mathcal{D}_k$ satisfies both (28) and (28), thus $\mathcal{D}_k \subseteq \mathcal{F}_k$. Therefore, if $P[\liminf_{k\to\infty} \mathcal{D}_k] = 1$, then (34) also holds.

In parts i) and ii) of the proof below, we show that conditioned on $\mathcal{D}_k$ (28) and (29) hold respectively. In part iii), we show almost sure occurrence of $\mathcal{D}_k$ as $k \to \infty$, thus concluding the proof.

1) Proof of Converse: For a given $\delta > 0$, let us assume that the event $\mathcal{D}_k$ happens. Also, let

$$2s \leq r_0. \tag{35}$$

In order to be able to apply Theorem 1, we need to determine the effective loss $1/G(\mathcal{R})$ of this network for $\mathcal{R} = \{2,\ldots,k\}$. To do so, we first upper bound the quantity $\sum_{j\in\mathcal{R}\setminus\{i\}} \mathbb{E}[|h_{ij}|^2] = \sum_{j\in\mathcal{R}\setminus\{i\}} g(r_{ij})$ for any node $i \in \{1,\ldots,k\}$; we show that the upper bound is proportional to k/A_k. Instead of directly evaluating the total channel gain from node i to all the nonsource nodes, we bound it from above by summing the maximum possible gains to all the nodes falling in a cell ℓ vertical, horizontal, or diagonal steps away, for $\ell = 0,1,\ldots,(\sqrt{A_k}/s)-1$. Thereafter, it is just a matter of simplifying the terms. Care needs to be exercised, treating those cells falling in the near-field separately from those in the far-field. Therefore, for any node $i \in \{1,\ldots,k\}$

$$\begin{aligned}
&\sum_{\substack{j=2\\ j\neq i}}^{k} g(r_{ij}) \\
&\leq \left((1+\delta)\frac{(k-1)s^2}{A_k}+1\right)\left[\bar{g} + \sum_{\ell=1}^{(\sqrt{A_k}/s)-1} 8\ell\, g((\ell-1)s)\right] && (36)\\
&\leq \left((1+\delta)\frac{(k-1)s^2}{A_k}+1\right) \\
&\quad\times\left[\left(1+8\sum_{\ell=1}^{\lceil r_0/s\rceil}\ell\right)\bar{g} + \frac{8}{s^\alpha}\sum_{\ell=\lceil r_0/s\rceil+1}^{(\sqrt{A_k}/s)-1}\frac{\ell}{(\ell-1)^\alpha}\right] && (37)\\
&\leq (1+2\delta)\frac{(k-1)s^2}{A_k}\left[\left(1+12\left(\frac{r_0}{s}\right)^2\right)\bar{g} + \frac{2^\alpha}{\alpha-2}\frac{3}{s^2 r_0^{\alpha-2}}\right] && (38)\\
&= (1+2\delta)\frac{(k-1)}{A_k}\left[(12.25)\,\bar{g}r_0^2 + \frac{2^\alpha}{\alpha-2}\frac{3}{r_0^{\alpha-2}}\right] && (39)\\
&\leq \frac{(k-1)\log_e 2}{c_1 A_k} && (40)
\end{aligned}$$

where

$$c_1 = \frac{4\log_e 2}{(1+2\delta)\left(49\,\bar{g}r_0^2 + \frac{2^{\alpha+2}}{\alpha-2}\frac{3}{r_0^{\alpha-2}}\right)}. \tag{41}$$

The explanation of the steps (36)–(40) is the following. Suppose that node i falls in cell C. Consider the set of cells exactly ℓ horizontal, vertical, or diagonal steps away from C (the only cell $\ell = 0$ steps away is C itself). There are at most $\max\{1, 8\ell\}$ cells ℓ steps away from C. The channel gain from node i to a node in any cell exactly ℓ steps away is not more than $g((\ell-1)s)$ for $\ell \geq 1$. For $\ell = 0$, the channel gain is not more than $\bar{g}$. Furthermore, since we assume that the event $\mathcal{D}_k$ occurs, the maximum number of nodes in a cell is less than $((1+\delta)(k-1)s^2/A_k)+1$. This gives us (36). For $\ell \geq \lceil r_0/s\rceil + 1$, since the minimum distance between the nodes is greater than r_0, the upper bound on gain is

$$g((\ell-1)s) \leq \frac{1}{s^\alpha(\ell-1)^\alpha}. \tag{42}$$

This immediately leads to (37). Step (38) requires a number of minor simplifications. First

$$(1+\delta)\frac{(k-1)s^2}{A_k} + 1 \leq (1+2\delta)\frac{(k-1)s^2}{A_k} \tag{43}$$

for all k large enough. Due to (35), the following inequalities hold:

$$\left\lceil\frac{r_0}{s}\right\rceil + 1 < 2\frac{r_0}{s} \tag{44}$$

$$\left\lceil\frac{r_0}{s}\right\rceil - 1 \geq \frac{1}{2}\frac{r_0}{s} \tag{45}$$

and

$$\left\lceil\frac{r_0}{s}\right\rceil < \frac{3}{2}\frac{r_0}{s}. \tag{46}$$

Thus, we have the following bound on the first sum in (37):

$$\sum_{\ell=1}^{\lceil r_0/s\rceil} \ell < \frac{3}{2}\left(\frac{r_0}{s}\right)^2 \tag{47}$$

since the sum of first n natural numbers is $\frac{1}{2}n(n+1)$. The other sum in (37) can be bounded as

$$\sum_{\ell=\lceil r_0/s\rceil+1}^{(\sqrt{A_k}/s)-1}\frac{\ell}{(\ell-1)^\alpha} = \sum_{\ell=\lceil r_0/s\rceil}^{(\sqrt{A_k}/s)-2}\frac{\ell+1}{\ell^\alpha} \tag{48}$$

$$\leq \frac{3}{2} \sum_{\ell=\lceil r_0/s \rceil}^{\infty} \frac{1}{\ell^{\alpha-1}} \tag{49}$$

$$\leq \frac{3}{2} \int_{\lceil r_0/s \rceil - 1}^{\infty} \frac{1}{u^{\alpha-1}}\, du \tag{50}$$

$$= \frac{3}{2(\alpha-2)} \frac{1}{\left(\lceil \frac{r_0}{s} \rceil - 1\right)^{\alpha-2}} \tag{51}$$

where we have used the fact that $x+1 \leq \frac{3}{2}x$ for all $x \geq 2$ in (49). Note that (51) is valid only for $\alpha > 2$.

Having bounded the total gain from any node i in (40), we get the following bound on effective network loss:

$$G(\mathcal{R}) \leq \frac{\log_e 2}{c_1 A_k} \tag{52}$$

which implies, by Theorem 1

$$\frac{E_b}{N_0}_{\min} \geq c_1 A_k \tag{53}$$

for all k large enough. Since the choice of $\delta > 0$ in (41) is arbitrary, we pick $\delta = 1/2$ to get the requisite lower bound on the minimum energy per bit for c_1 given by (30).

Proof of Achievability: For any $\epsilon_1, \epsilon_2 > 0$, we show that, conditioned on $\mathcal{D}_k$

$$\texttt{FLOOD}\left(\frac{N_0 \log_e 2}{g(\sqrt{8}s)} + \epsilon_1, \frac{(1+\epsilon_2)A_k N_0 \log_e 2}{s^2(k-1)g(\sqrt{8}s)}\right) \tag{54}$$

manages to reliably communicate the common message to all the nodes.

Set $s = r_0/\sqrt{8}$ so that we can replace $g(\sqrt{8}s)$ with $r_0^{-\alpha}$. Thus, the total energy consumption per bit of (54) is

$$E_{b\text{total}} \leq \left(r_0^{\alpha} + (1+\epsilon_2)8A_k r_0^{\alpha-2}\right) N_0 \log_e 2 + \epsilon_1 \tag{55}$$

$$\leq \left((2+\epsilon_2)8 r_0^{\alpha-2} \log_e 2\right) A_k N_0 + \epsilon_1 \tag{56}$$

where we have used the lower bound (27) on A_k in simplifying (55) to (56).

Our next step is to show that the scheme in (54) is able to reach all the nodes. First, for the given value of ϵ_2, we choose any

$$0 < \delta < \frac{\epsilon_2}{2(1+\epsilon_2)}$$

in the definition of event $\mathcal{D}_k$. Thus, if $\mathcal{D}_k$ occurs, then all the cells have at least

$$\frac{1+\frac{\epsilon_2}{2}}{1+\epsilon_2} \frac{(k-1)r_0^2}{8A_k} \tag{57}$$

nodes.

Let $T_k = \sqrt{8A_k}/r_0$ be the maximum number of time slots in the flooding scheme. Suppose that a nonsource node i belongs to cell C. For any cell C, there is a sequence $(C_1, C_2, \ldots, C_T)$ of horizontally, vertically, or diagonally adjacent cells such that C_1 is the origin cell and $C_T = C$, for some $T \leq T_k$. We now present an argument to show that the nodes in cell C_t successfully decode by the end of slot $t-1$.

In the first time slot, the source node transmits with energy per bit

$$E_{b_1} > r_0^{\alpha} N_0 \log_e 2 \tag{58}$$

which implies that the received energy per bit at all the nodes within a radius of r_0 (which includes all the nodes falling within two cells of the origin) is strictly greater than $N_0 \log_e 2$. Therefore, all the nodes in the cell C_2 are able to decode the message reliably with vanishing probability of error [25, Th. 1]. Now, suppose that all the nodes in cell C_{t+1} are able to decode without error by the end of slot t. This implies that all the nodes in the cell C_{t+1} would have transmitted the correct message by the end of slot $t+1$ (possibly in different slots). The energy per bit of their transmissions is

$$E_{b_2} = \frac{(1+\epsilon_2)8A_k r_0^{\alpha-2} N_0 \log_e 2}{k-1}. \tag{59}$$

For any node i in the cell C_{t+2}, the gain from any of the nodes in C_{t+1} is at least $r_0^{-\alpha}$ since the distance between them is at most r_0. Furthermore, the minimum number of nodes in the cell C_{t+1} is given by (57). Therefore, the total received energy at node i is

$$E_b^r \geq r_0^{-\alpha}\left(\frac{1+\frac{\epsilon_2}{2}}{1+\epsilon_2}\frac{(k-1)r_0^2}{8A_k}\right) \times \left(\frac{(1+\epsilon_2)8A_k r_0^{\alpha-2} N_0 \log_e 2}{k-1}\right) \tag{60}$$

$$> N_0 \log_e 2. \tag{61}$$

Moreover, all this energy is received by the end of slot $t+1$. Therefore, the node i (and hence, every node in C_{t+2}) successfully decodes the message by the end of slot $t+1$ with vanishing probability of error. Thus, by inductive reasoning, the flooding scheme is able to reach out to every cell C (and hence, to every node) by the end of slot $T_k - 1$.

Note that since the choice of ϵ_1 and ϵ_2 is arbitrary, from (56), we get that $E_{b\text{flood}}$ satisfies the condition laid out in the statement of Theorem 2 with the constant c_2 given by (31), for $\epsilon_1 \to 0$ and $\epsilon_2 = 1$.

2) Almost Sure Occurrence of $\mathcal{F}_k$: Note that the condition (26) implies that $A_k \leq k$ for all k large enough, and that for any constant $c' \geq 0$, the value of $(k-1)/A_k$ is greater than $c' \log_e k$ for all k large enough. Also, the bound in (53) holds for all k large enough. Pick a k' large enough that these three conditions are satisfied (with $c' = 6/(\delta^2(1-\delta)s^2)$) for any $k \geq k'$. Now, consider the sum

$$\sum_{k=2}^{\infty} P[\mathcal{D}_k^c] \leq \frac{2}{s^2} \sum_{k=2}^{\infty} A_k \exp\left(-\delta^2(1-\delta)\frac{(k-1)s^2}{2A_k}\right) \tag{62}$$

$$\leq c + \frac{2}{s^2} \sum_{k=k'}^{\infty} k \exp\left(-\delta^2(1-\delta)\frac{(k-1)s^2}{2A_k}\right) \tag{63}$$

$$\leq c + \frac{2}{s^2} \sum_{k=k'}^{\infty} k \exp\left(-3 \log_e k\right) \tag{64}$$

$$= c + \frac{2}{s^2}\sum_{k=k'}^{\infty}\frac{1}{k^2} \quad (65)$$

$$\leq c + \frac{2}{s^2}\frac{\pi^2}{6} \quad (66)$$

$$< \infty \quad (67)$$

where c is some real positive constant. Inequality (62) is due to Lemma 2; and the inequalities (63) and (64) are due to the choice of k'.

Therefore, since $\sum_{k=2}^{\infty} P[\mathcal{D}_k^c]$ is finite, by the Borel–Cantelli lemma, we conclude that the event $\mathcal{D}_k^c$ occurs infinitely often with probability 0. Since $\mathcal{F}_k^c \subseteq \mathcal{D}_k^c$, the event $\mathcal{F}_k^c$ occurs infinitely often also with probability 0. Hence, with probability 1, the event $\mathcal{F}_k$ happens for all but a finite number of k, as $k \to \infty$. □

Remark 4: Note that Theorem 2 also includes the case of constant area, i.e., $A_k = A$. In this case, the minimum energy per bit is $\Theta(1)$ even as the number of nodes becomes large. To see this simply, note that even if two nodes are very close, the path loss model dictates that there be at least a constant energy consumption per bit $(N_0 \log_e 2/\bar{g})$ in communicating information from one node to the other. On the other hand, a single-shot transmission by the source also has a constant achievable energy per bit $(N_0 \log_e 2/g(\sqrt{2A}))$.

However, if the area A_k grows with k, the single-shot transmission scheme requires energy proportional to ${A_k}^{\alpha/2}$. This is improved upon in Theorem 2 by multihop-based flooding. On the other hand, the advantages of single-shot transmission include small delay (just a single time slot) which does not grow with the network size, and its resilience to inhomogeneity in the node distribution.

Remark 5: Theorem 2 states that if the node density is guaranteed to increase at a large enough rate [i.e., density is $\omega(\log k)$], the total energy consumption per bit depends only on the area and is independent of the number of nodes. Intuitively, once the spacing between nodes gets small enough, the multihop advantage vanishes since the nearest neighbors are now in the near-field. Thus, reducing the transmission radius does not improve the efficiency any further, and the energy expenditure is proportional to the area that needs to be covered. At the same time, increasing node density is critical since it ensures the presence of enough neighbors even when the nodes are randomly placed. We discuss the important issue of deviations in node placement again in Section V-C (for extended networks) and in Section VI.

Remark 6: If $\alpha = 2$, the sum on the left-hand side of (48) would grow as $\Theta(\log A_k)$ which weakens (40) to a constant times $(k-1)\log A_k/A_k$. Therefore, the bound now becomes

$$\frac{E_b}{N_0}_{\min} \geq c_1' \frac{A_k}{\log A_k} \quad (68)$$

for some constant $c_1' > 0$. This was shown to be achievable in [19] in a different setting (as mentioned in Section I-A).

Remark 7: Note that increasing the number of nodes in the network increases the net available energy in the network. On the other hand, from Theorem 2, there is no substantial increase in the minimum energy per bit for broadcasting to these additional nodes. Thus, a simple way of decreasing the energy burden (per bit) on the individual nodes is to increase their number, which would also increase the network lifetime.

C. Random Extended Networks

The extended random network case differs from the dense case because the density of the nodes is now a constant

$$\lambda = \frac{k}{A_k} \quad (69)$$

in nodes per square meter. As in dense networks, the source node is placed at the origin and the rest $k-1$ nonsource nodes are distributed independently and uniformly over the area A_k.

This case is different than the dense networks case since the node density is now fixed. This limits the advantage that can be had by nearest neighbor hopping since the variations in node placement could lead to isolated nodes such that their nearest neighbor distance grows as $\log k$ as $k \to \infty$. Note that in contrast to extended networks, the nearest neighbor hopping advantage in the dense networks is limited due to the path loss model.

Theorem 3: With probability 1, the node placement is such that

$$c_1 \leq \frac{1}{k}\frac{E_b}{N_0}_{\min} \quad (70)$$

and

$$\frac{1}{k(\log_e k)^{\alpha/2}}\frac{E_b}{N_0}_{\text{flood}} \leq c_2 \quad (71)$$

for all but a finite number of k, where

$$c_1 = \begin{cases} \dfrac{\log_e 2}{2^4 3^{\alpha+2} e\zeta(\alpha-1)}\lambda^{-\alpha/2}, & \text{for } \lambda < \dfrac{1}{9r_0^2} \\ \dfrac{\log_e 2}{2^5 3^3\left(\bar{g}r_0^2 + \dfrac{1}{(\alpha-2)6^{\alpha} r_0^{\alpha-2}}\right)}\lambda^{-1}, & \text{for } \lambda \geq \dfrac{1}{9r_0^2} \end{cases} \quad (72)$$

where $\zeta(\cdot)$ is the Riemann zeta function. And

$$c_2 = 3\,2^{2\alpha-1}(\log_e 2)\lambda^{-\alpha/2}. \quad (73)$$

Proof: Let $\mathcal{F}_k$ be the collection of all the node placement realizations for k nodes for which both (70) and (71) hold. To prove Theorem 3, we only need to show that

$$P[\liminf_{k\to\infty} \mathcal{F}_k] = 1. \quad (74)$$

To do so, in parts i) and ii) of the proof, we construct events $\mathcal{D}_k$ and $\mathcal{E}_k$, respectively, such that, conditioned on the events, the bounds (70) and (71) hold, respectively. Using results on the almost sure occurrence of $\mathcal{D}_k$ and $\mathcal{E}_k$, in part iii), we conclude that $\mathcal{F}_k$ occurs almost surely, as $k \to \infty$.

1) Proof of Converse: For simplicity, we assume that the number of nodes k is a square integer larger than 1. Partition the network area into square cells with side length $\lambda^{-1/2}$. This implies that there are k cells in the network. Let $\mathcal{C}$ be the set of cells, and let $\nu(C)$ denote the number of nodes in cell $C \in \mathcal{C}$. Next, right at the center of each cell, consider a small square *window* of side length $\beta\lambda^{-1/2}$, where $0 \leq \beta \leq 1$ is a constant to

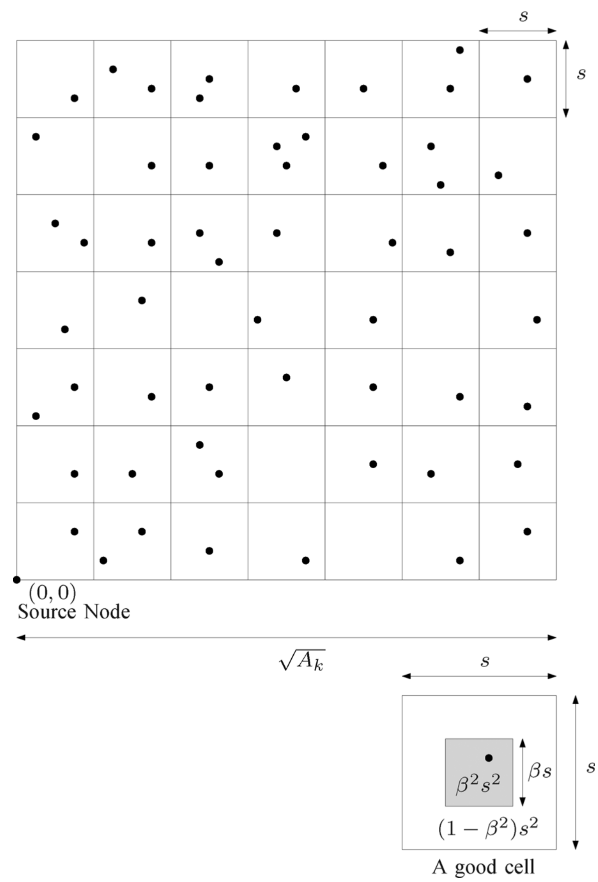

Fig. 3. Random extended network.

be selected later. Define a nonorigin cell to be *good* if it contains exactly one node within its window and no nodes outside the window. The nonsource node falling in a good cell is called a *good node*. See Fig. 3. Let the set of good nodes be $\mathcal{R}_1 \subset \{2, \ldots, k\}$. The number of good nodes (cells) is denoted by k_1. The set $\mathcal{R}_1$ is our destination set, for which we will bound the effective network loss $1/G(\mathcal{R}_1)$.

For any $\delta > 0$, define a *good placement* event $\mathcal{D}_k$ as the collection of node placements for which

$$k_1 \geq (1-\delta)\beta^2 \frac{(k-1)^k}{k^{k-1}}. \tag{75}$$

The next result lower bounds the probability of a good placement.

Lemma 3:

$$P[\mathcal{D}_k^c] \leq 2\exp\left(-\frac{1}{2}\delta^2\beta^4 \frac{(k-1)^{2k-1}}{k^{2k-2}}\right) \tag{76}$$

for all $k \geq 2$.

Proof: See Appendix III. □

Let us assume, for the time being, that the event $\mathcal{D}_k$ happens and focus our attention on finding an upper bound on $G(\mathcal{R}_1)$. For any node $i \in \{1, \ldots, k\}$, we have the following upper bound on total gain from node i to the nodes in $\mathcal{R}_1$:

$$\sum_{j \in \mathcal{R}_1 \setminus \{i\}} g(r_{ij}) \leq \sum_{\ell=1}^{\sqrt{k}-1} 8\ell\, g(r_{ij}) \tag{77}$$

$$\leq \sum_{\ell=1}^{\sqrt{k}-1} 8\ell\, g\left(\frac{(\ell-1)}{\sqrt{\lambda}} + \frac{1-\beta}{2\sqrt{\lambda}}\right) \tag{78}$$

where (77) and (78) are obtained as follows. Consider all those good cells lying exactly ℓ horizontal, vertical, or diagonal steps away from the cell containing node i. Since there are at most 8ℓ cells ℓ steps away, there are at most as many good cells ℓ steps away, implying (77). Furthermore, a good node ℓ steps away is at a distance of $((\ell-1)/\sqrt{\lambda}) + ((1-\beta)/2\sqrt{\lambda})$ or greater. Thus, the gain to that cell cannot exceed $g(((l-1)/\sqrt{\lambda}) + ((1-\beta)/2\sqrt{\lambda}))$, implying (78).

To further simplify the right-hand side of (78), we need to consider the following two cases for λ.

1) $\lambda < 1/(9r_0^2)$. Let $\beta = 1/3$. Continuing from (78), an upper bound on the total gain from node i to the nodes in $\mathcal{R}_1$ is

$$\sum_{j \in \mathcal{R}_1 \setminus \{i\}} g(r_{ij}) \leq \sum_{\ell=1}^{\infty} 8\ell\, g\left(\frac{1-\beta}{2\sqrt{\lambda}}\ell\right) \tag{79}$$

$$\leq 8\left(\frac{1-\beta}{2\sqrt{\lambda}}\right)^{-\alpha} \sum_{\ell=1}^{\infty} \ell^{-(\alpha-1)} \tag{80}$$

$$= 2^3\, 3^\alpha \zeta(\alpha-1)\lambda^{\alpha/2} \tag{81}$$

where, in (81), $\zeta(\cdot)$ is the Riemann zeta function which is finite for all real arguments greater than 1; (79) is obtained from (78) by observing that

$$((\ell-1)/\sqrt{\lambda}) + ((1-\beta)/2\sqrt{\lambda}) \geq (1-\beta)\ell/2\sqrt{\lambda} \tag{82}$$

for all $\ell \geq 1$. Also, $(1-\beta)/2\sqrt{\lambda} > r_0$ for the given value of β, due to the defining condition of this case. Therefore, all the gain terms are given by the far-field case (24), which immediately implies (80).

Since the bound in (81) holds for every node $i \in \{1, \ldots, k\}$, from (75) and (81), we get the following bound on effective network loss:

$$G(\mathcal{R}_1) \leq \frac{2^3\, 3^\alpha \zeta(\alpha-1)\lambda^{\alpha/2}}{(1-\delta)\beta^2 \dfrac{(k-1)^k}{k^{k-1}}} \tag{83}$$

$$\leq \frac{2^3\, 3^{\alpha+2} e \zeta(\alpha-1)\lambda^{\alpha/2}}{(1-2\delta)k} \tag{84}$$

for all k large enough. Inequality (84) is due to the fact that since $\lim_{k\to\infty}(1-1/k)^k = 1/e$, for any $\delta > 0$

$$\left(1-\frac{1}{k}\right)^k \geq \left(\frac{1-2\delta}{1-\delta}\right)\frac{1}{e} \tag{85}$$

for all k large enough.

Using (84) in Theorem 1 immediately gives us that for all k large enough

$$\frac{E_b}{N_{0\,\min}} \geq (1-2\delta)\left(\frac{\log_e 2}{2^3 3^{\alpha+2} e\zeta(\alpha-1)}\right) k\lambda^{-\alpha/2}. \quad (86)$$

Since the choice of $\delta > 0$ is arbitrary, taking $\delta = 1/4$, we obtain the claimed bound for $\lambda < 1/(9r_0^2)$.

2) $\lambda \geq 1/(9r_0^2)$. Now, let $\beta = 1$.
Define

$$L \triangleq \lceil 6r_0\sqrt{\lambda}\rceil \quad (87)$$

which is roughly the number of cells beyond which the far-field model is valid. Note that $L \geq 2$. Continuing from (78), an upper bound on the total gain from node i to the nodes in $\mathcal{R}_1$ is

$$\begin{aligned}\sum_{j\in\mathcal{R}_1\setminus\{i\}} g(r_{ij}) &\leq \sum_{\ell=1}^{L} 8\ell\,\bar{g} + \sum_{\ell=L+1}^{\sqrt{k}-1} 8\ell\, g\left(\frac{(\ell-1)}{\sqrt{\lambda}}\right) && (88)\\ &\leq 8\bar{g}\sum_{\ell=1}^{L}\ell + 8\lambda^{\alpha/2}\sum_{\ell=L+1}^{\infty}\frac{\ell}{(\ell-1)^{\alpha}} && (89)\\ &\leq 4L(L+1)\bar{g} + 12\lambda^{\alpha/2}\sum_{\ell=L}^{\infty}\ell^{-(\alpha-1)} && (90)\\ &\leq 12\left(6r_0\sqrt{\lambda}\right)^2\bar{g} + \frac{12\lambda^{\alpha/2}}{\alpha-2}\,\frac{\lambda^{1-(\alpha/2)}}{6^{\alpha-2}r_0^{\alpha-2}} && (91)\\ &= 2^4\,3^3\left(r_0^2\bar{g} + \frac{1}{(\alpha-2)2^{\alpha}3^{\alpha}r_0^{\alpha-2}}\right)\lambda. && (92)\end{aligned}$$

Inequality (88) is obtained from (78) by breaking the sum into two parts, and bounding the gain terms in the first sum by $\bar{g}$ and in the second sum by $g((\ell-1)/\sqrt{\lambda})$; (89) is due to the fact that for $\ell \geq L+1$, the distance $(\ell-1)/\sqrt{\lambda}$ is greater than r_0; the explanation for (90) and (91) is similar to that of (36)–(40).

Since (92) is valid for every node $i \in \{1,\ldots,k\}$, we can bound $G(\mathcal{R}_1)$ and thus find the lower bound for this case in a manner similar to the previous case.

To conclude this part of the proof, we show that the converse bound is violated (i.e., $\mathcal{D}_k^c$ occurs) infinitely often with probability 0. To do so, consider the sum

$$\begin{aligned}\sum_{k=2}^{\infty} P[\mathcal{D}_k^c] &\leq 2\sum_{k=2}^{\infty}\exp\left(-\frac{1}{2}\delta^2\beta^4\frac{(k-1)^{2k-1}}{k^{2k-2}}\right) && (93)\\ &\leq 2\sum_{k=2}^{\infty}\left(\exp\left(-\frac{\delta^2}{2\,3^4}\left(1-\frac{1}{k}\right)^{2k-2}\right)\right)^{k-1} && (94)\\ &\leq 2\sum_{k=2}^{\infty}c^{k-1} && (95)\\ &< \infty && (96)\end{aligned}$$

where (93) is from Lemma 3; inequality (94) is valid for both $\beta = 1/3$ and 1 (which covers both cases); the bound in (95) is obtained by noticing that $(1-1/k)^{2(k-1)}$ is greater than $1/e^2$ for all $k \geq 2$. The value of c in (95) is given by

$$c = \exp\left(-\frac{\delta^2}{2\,3^4\,e^2}\right). \quad (97)$$

Since $c < 1$, the infinite geometric progression in (95) converges to a finite value. In view of the summability of $P[\mathcal{D}_k^c]$, we invoke the Borel–Cantelli lemma to conclude that

$$P[\limsup_{k\to\infty}\mathcal{D}_k^c] = 0. \quad (98)$$

2) Proof of Achievability: For any $\epsilon > 0$, consider the scheme

$$\texttt{FLOOD}\left(\frac{N_0\log_e 2}{g(\sqrt{8}s_k)} + \frac{\epsilon}{k}, \frac{N_0\log_e 2}{g(\sqrt{8}s_k)} + \frac{\epsilon}{k}\right) \quad (99)$$

where $s_k > 0$ (to be selected later) is the side length of the cells in the network. For the sake of simplicity, $\sqrt{A_k}$ is assumed to be a multiple of s_k. Denote by $\mathcal{E}_k$ the event that no cell is empty. Let us assume, for the time being, that $\mathcal{E}_k$ occurs. Since all the cells are nonempty, for every cell C, we have a sequence (of length at most $\sqrt{A_k}/s_k$) of adjacent (horizontally, vertically, or diagonally) nonempty cells which begins at the cell containing the origin and terminates at cell C. Therefore, there is a path of nodes from the source node to any other node such that two consecutive nodes are within a distance of $\sqrt{8}s_k$ of each other. This implies that if a node that has already decoded a message transmits with energy per bit greater than $N_0\log_e 2/g(\sqrt{8}s_k)$, its transmission will be received by nodes in the neighboring cells with sufficient energy to decode the message. Thus, the multihop scheme (99) suffices to reach every node.

For any $\delta > 0$, set

$$s_k^2 = \frac{(2+\delta)\log_e A_k}{\lambda}. \quad (100)$$

Since s_k grows unbounded with k, $\sqrt{8}s_k \geq r_0$ eventually. Therefore, we can replace $g(\sqrt{8}s_k)$ with

$$\frac{1}{\left(\sqrt{8}s_k\right)^{\alpha}} = \frac{\lambda^{\alpha/2}}{(8(2+\delta)\log_e A_k)^{\alpha/2}} \quad (101)$$

for all k large enough. Thus, from (22), the $E_{b\text{total}}$ of this algorithm satisfies

$$E_{b\text{total}} \leq \frac{(8(2+\delta)\log_e A_k)^{\alpha/2} k N_0\log_e 2}{\lambda^{\alpha/2}} + \epsilon. \quad (102)$$

Substituting the value of A_k from (69) and noting that the choices of ϵ and δ are arbitrary, we immediately get that

$$\frac{E_b}{N_{0\,\text{flood}}} \leq c_2\, k\lambda^{-\alpha/2}(\log_e k)^{\alpha/2} \quad (103)$$

for all k large enough and c_2 as given in (73).

Hence, for all k large enough, every node placement in $\mathcal{E}_k$ satisfies the bound (71). We now show that the converse bound is violated (i.e., $\mathcal{E}_k^c$ occurs) infinitely often with probability 0.

For every k, let C_1 be any fixed nonsource cell. Consider the sum

$$\sum_{k=2}^{\infty} P[\mathcal{E}_k^c] = \sum_{k=2}^{\infty} P[\exists C \in \mathcal{C} : \nu(C) = 0] \quad (104)$$

$$\leq \sum_{k=2}^{\infty} \left(\frac{A_k}{s_k^2} - 1\right) P[\nu(C_1) = 0] \quad (105)$$

$$\leq \sum_{k=2}^{\infty} \frac{A_k}{s_k^2} \left(1 - \frac{s_k^2}{A_k}\right)^{k-1} \quad (106)$$

$$\leq \sum_{k=2}^{\infty} \frac{k}{(2+\delta)\log A_k} \exp\left(-(2+\delta)\frac{k-1}{k}\log A_k\right) \quad (107)$$

$$\leq c + \sum_{k=k_1}^{\infty} \frac{k}{(2+\delta)\log A_k} \frac{1}{{A_k}^{2+\delta/2}} \quad (108)$$

$$< \infty \quad (109)$$

where (105) is from the union bound over all the nonorigin cells in the network; (106) is obtained by noting that the probability of the nonorigin cell C_1 being empty is $\left(1 - (s_k^2/A_k)\right)^{k-1}$; noting that $1 - x \leq \exp(-x)$ for all $x \geq 0$ and together from (69) and (100), we obtain (107); since for any given $\delta > 0$

$$(2+\delta)\frac{k-1}{k} \geq 2 + \frac{\delta}{2} \quad (110)$$

for all $k \geq k_1$ for all large enough k_1, (107) simplifies to (108), where $c < \infty$ is some constant; the series in (108) converges since A_k is linear in k, and $\sum_{k=1}^{\infty} k^{-(1+\delta_1)}$ converges for any $\delta_1 > 0$. Therefore, by the Borel–Cantelli lemma, from (109), we get that

$$P[\limsup_{k\to\infty} \mathcal{E}_k^c] = 0. \quad (111)$$

3) Almost Sure Occurrence of $\mathcal{F}_k$: Note that $\mathcal{D}_k \cap \mathcal{E}_k \subseteq \mathcal{F}_k$, which implies $\mathcal{F}_k^c \subseteq \mathcal{D}_k^c \cup \mathcal{E}_k^c$. Therefore, (98) and (111) imply that $P[\limsup_{k\to\infty} \mathcal{F}_k^c] = 0$ which is the same as $P[\liminf_{k\to\infty} \mathcal{F}_k] = 1$. Hence, with probability 1, the event $\mathcal{F}_k$ happens for all but a finite number of k, as $k \to \infty$. □

Remark 8: The result in Theorem 3 shows that the gap between the upper and lower bounds is of the order $O\left((\log k)^{\alpha/2}\right)$ which is small compared to the order of growth $\Omega(k)$ of the lower bound on minimum energy per bit. This gap is due to the sparse nature of the extended networks which would cause an occasional node to be isolated (i.e., have a nearest neighbor distance growing as $\Theta(\log k)$). In order to reach such nodes, uniform power allocation requires the energy per bit quota to be uniformly scaled by an additional factor of $\Omega\left((\log k)^{\alpha/2}\right)$.

Note that the requirement on the uniform energy allocation is imposed by the minimal information setting where we assume that the nodes have no information about their neighbors. If the nodes do have the nearest neighbor information, the gap between the upper and lower bounds could potentially be reduced. Optimal energy allocation can then be done, along the lines of the optimal allocation suggested in [11], [15], and [19] (see the discussion in Section I-A).

Remark 9: According to Theorem 3, the energy per bit requirement in extended networks is approximately linearly proportional to the number of nodes. This implies that there is no advantage in increasing the number of nodes in the networks (in contrast to the dense networks, Remark 7). Therefore, each node adds about a constant [c_1, as given by (72)] requirement to the energy per bit. This requirement can also be viewed as the price paid by each node to be a part of the broadcast network, and can be justified by noticing that each node is extending the reach of the network by some constant area.

Remark 10: Note that the results in Theorem 3 are in terms of the total number of nodes k. Since the area of the network A_k is proportional to k, the inequalities (70) and (71) still hold with k replaced by A_k and different constants c_1 and c_2. Note that while the lower bound has the same growth with A_k as for the dense networks, the upper bound is weaker than the upper bound in dense networks (see Remark 8). Furthermore, the behavior of the constants differs for the two cases of the networks. The proof techniques for Theorem 3 also differ from those of Theorem 2, for both upper and lower bounds.

VI. Regular Networks

In both dense and extended random networks, we saw that the proposed bounds on $E_{b\min}$ hold almost surely as $k \to \infty$. These bounds fail to hold when there is a nonfavorable placement of nodes, the probability of which is nonzero when k is finite. In this section, we consider finite networks where there is some regularity in node placement. In Theorem 4, we show that regularity does not only give deterministic results for finite sized networks, but the upper and lower bounds are tight up to a constant factor.

In regular networks, the network area is divided into square cells with side length s. Each cell has a square window in its center. The window is assumed to occupy a fraction $0 \leq \beta^2 < 1$ of the cell area. The regularity condition is that each cell contains exactly one node in its window and no other node outside the window. Each node can be arbitrarily placed within its window. Note that the number of cells is the same as the number of nodes k. The source node lies in the window of the origin cell. See Fig. 4. For the sake of simplicity, we assume $\sqrt{k}$ is an integer larger than 1. For $x, y = 0, 1, \ldots, \sqrt{k} - 1$, the notation $C(x, y)$ is used to denote the cell with the lower left corner on the coordinates (xs, ys). As discussed, the node in cell $C(x, y)$ lies within the square having its diagonal coordinates at $(xs + ((1-\beta)s/2), ys + ((1-\beta)s/2))$ and $(xs + ((1+\beta)s/2), ys + ((1+\beta)s/2))$. We retain the path loss model of Section V-A.

The parameter β denotes the flexibility in the placement of the nodes. $\beta = 0$ is the case when there is no flexibility and all the nodes fall exactly on the lattice points.

The following result provides an upper bound on the ratio $E_{b\text{flood}}/E_{b\min}$ that is independent of the number of nodes and the cell size.

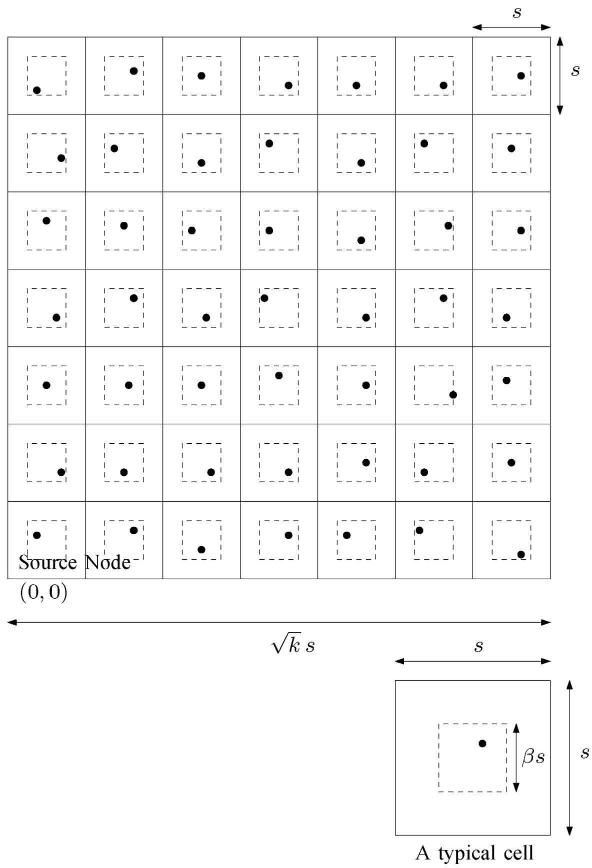

Fig. 4. Regular network.

Theorem 4: For any regular network with $k \geq 2$ nodes and cell size $s > 0$

$$\frac{E_{b\text{flood}}}{E_{b\min}} \leq c_1 \tag{112}$$

where

$$c_1 \leq \max\left\{\frac{2^{\frac{3\alpha}{2}+4}\zeta(\alpha-1)}{(1-\beta)^\alpha}, (2r_0)^\alpha \bar{g}, \frac{2^{\frac{3\alpha}{2}+3}r_0^\alpha}{(1-\beta)^{\alpha+2}}\left(\bar{g}+\frac{2^{\alpha-2}}{(\alpha-2)r_0^\alpha}\right)\left(1+4(1-\beta)^2\right)\right\}. \tag{113}$$

Proof: The proof is similar to the proofs of Theorems 2 and 3, and is provided in Appendix IV. □

Remark 11: Theorem 4 implies that regardless of the density of nodes or the area of the network, if the nodes are spread out "regularly" over the network area, the FLOOD algorithm has an achievable energy per bit within a constant factor of the lower bound. This justifies our intuition that the gap between the upper and lower bounds in extended networks is due to variations (from a regular placement) in the node placements.

VII. Discussion and Conclusion

Motivated by energy limited wireless network scenarios, we have studied the minimization of the energy per bit required for broadcasting.

Using information-theoretic tools, we have first established a lower bound on the minimum energy per bit for reliable multicasting in arbitrary wireless networks. The network is assumed to operate in the wideband regime, and the communication channel is affected by circularly symmetric fading (not known at the transmitters) and additive white Gaussian noise. The lower bound is based on a fundamental quantity—effective network loss—which depends only on the expected power attenuations between pairs of nodes. This lower bound holds regardless of the available bandwidth, or the communication or cooperation schemes. It is convenient that the bound depends on the physical layer description of the network only through one scalar, the effective network loss.

Next, for the achievability part, we restrict the network information available at the nodes since such information is neither easy nor cheap (in terms of energy) to obtain in many practical scenarios. We show that for many classes of networks of practical interest, the lower bound can be approached by a simple flooding scheme even with minimal network information. The proposed flooding algorithm is a wideband, multistage, decode-and-forward communication scheme which can be implemented using simple repetition codes.

Using the proposed flooding algorithm, we are able to show that the minimum energy per bit is proportional to the area of the network for the case of dense random networks. For extended random networks which have constant node density, the minimum energy per bit is roughly linearly proportional to the number of nodes, with the gap between the upper and lower bound being a factor poly-logarithmic in the number of nodes. The bounds for both dense and extended networks hold almost surely as the network size grows large. The lower bounds on the minimum energy per bit are obtained assuming availability of the channel state information (at the receivers only) and the node location information. In contrast, the upper bounds are found by analyzing the flooding algorithm which employs uniform power allocation at the nonsource nodes. Therefore, as far as the order of scaling is concerned, knowledge about node locations or channel conditions (at the receivers) does not buy much in large random networks.

In order to understand the dependence of energy efficiency on the location of nodes, we have also studied finite-sized regular networks where nodes are restricted to their individual cells. We show that such regularity in node placement leads to upper and lower bounds that are within a constant factor of each other, regardless of the number of nodes and geographical size of the network.

Tightening the bounds in the nonasymptotic regime appears to be quite challenging in view of the fact that the minimum energy per bit remains unknown even for such basic building blocks as the relay channel.

Appendix I
Proof of Lemma 1

1) Proof: Consider a $(n, M, E_{\text{total}}, \epsilon)$ code over n channel uses. Let $E_{i,t}$ be the expected energy consumption of node i at time t, for $i = 1, \ldots, k$ and $t = 1, \ldots, n$. Therefore, the total energy consumption at node i is

$$E_i = \sum_{t=1}^{n} E_{i,t} \quad (114)$$

which satisfy

$$\sum_{i=1}^{k} E_i = E_{\text{total}}. \quad (115)$$

For the rest of the proof, $\mathbf{x}^{(n)} = (x_1^{(n)}, \ldots, x_k^{(n)})$ denotes the set of transmissions at all the nodes, and $\mathbf{x}_t = (x_{1,t}, x_{2,t}, \ldots, x_{k,t})$ is the set of symbols transmitted by all the nodes at time t.

For each node $j \in \mathcal{R}$, we derive a form of *cut-set* or *max-flow min-cut* bound (see [5, Th. 4] and [6, Th. 15.10.1]) in the following steps:

$$\begin{aligned}
&(1-\epsilon)\log_2 M \\
&\quad \le I(x_1^{(n)}; y_j^{(n)}) + 1 && (116)\\
&\quad \le I(\mathbf{x}^{(n)}; y_j^{(n)}) + 1 && (117)\\
&\quad = \sum_{t=1}^{n} I(\mathbf{x}^{(n)}; y_{j,t} | y_{j,1}, \ldots, y_{j,t-1}) + 1 && (118)\\
&\quad = \sum_{t=1}^{n} I(\mathbf{x}_t; y_{j,t} | y_{j,1}, \ldots, y_{j,t-1}) + 1 && (119)\\
&\quad \le \sum_{t=1}^{n} I(\mathbf{x}_t; y_{j,t}) + 1 && (120)\\
&\quad \le \sum_{t=1}^{n} I(\mathbf{x}_t; y_{j,t} | \mathbf{h}_{j,t}) + 1 && (121)\\
&\quad \le \sum_{t=1}^{n} \sup_{\substack{P_{\mathbf{x}_t}: \\ \mathbb{E}[|x_{i,t}|^2] \le E_{i,t} \text{for } i=1,\ldots,k}} I(\mathbf{x}_t; y_{j,t} | \mathbf{h}_{j,t}) + 1 && (122)\\
&\quad \le n \sup_{\substack{P_{\mathbf{x}}: \\ \mathbb{E}[|x_i|^2] \le E_i/n}} I(\mathbf{x}; y_j | \mathbf{h}_j) + 1 && (123)
\end{aligned}$$

where the explanation of steps (116)–(123) is the following. Inequality (116) is due to Fano's inequality. Inequality (117) is by expanding the set of random variables to $\mathbf{x}^{(n)}$. Applying the chain rule for mutual information to (117) gives us (118). Step (119) follows from the fact that $y_{j,t}$ depends on $\mathbf{x}^{(n)}$ only through the current transmissions $\mathbf{x}_t$. Similarly, since $(y_{j,1}, \ldots, y_{j,t-1})$—$\mathbf{x}_t$—$y_{j,t}$ form a Markov chain, (120) is also true. The random variables $\mathbf{x}_t$ and $\mathbf{h}_{j,t} = (h_{1j,t}, \ldots, h_{(j-1)j,t}, 0, h_{(j+1)j,t}, \ldots, h_{kj,t})^T$ are independent of each other, justifying (121). Applying the energy restriction $E_{i,t}$ on the symbol $x_{i,t}$ gives us (122). Finally, (123) is due to the concavity of the mutual information in the cost (in this case, power). Observe that if the supremum of the mutual information in the right-hand side of (123) is zero, then the mutual information term in the right-hand side of (116) is also zero which means that no reliable communication (i.e., $\epsilon \to 0$) is possible for large message sets (i.e., when $\log_2 M > 1$).

Since inequality (123) is valid for all nodes $j \in \mathcal{R}$, we can write

$$(1-\epsilon)\log_2 M \le \min_{j \in \mathcal{R}} n \sup_{\substack{P_{\mathbf{x}}: \\ \mathbb{E}[|x_i|^2] \le E_i/n}} I(\mathbf{x}; y_j | \mathbf{h}_j) + 1. \quad (124)$$

Therefore, the energy per bit of the code is as given by

$$\begin{aligned}
\frac{E_{\text{total}}}{\log_2 M} &\ge \left(1 - \epsilon - \frac{1}{\log_2 M}\right) \\
&\quad \times \frac{E_{\text{total}}}{n \min_{j \in \mathcal{R}} \sup_{\substack{P_{\mathbf{x}}: \\ \mathbb{E}[|x_i|^2] \le E_i/n}} I(\mathbf{x}; y_j | \mathbf{h}_j)} && (125)\\
&= \left(1 - \epsilon - \frac{1}{\log_2 M}\right) \\
&\quad \times \left(\frac{\min_{j \in \mathcal{R}} \sup_{\substack{P_{\mathbf{x}}: \\ \mathbb{E}[|x_i|^2] \le E_i/n}} I(\mathbf{x}; y_j | \mathbf{h}_j)}{\sum_{i=1}^{k} E_i/n} \right)^{-1} && (126)\\
&\ge \left(1 - \epsilon - \frac{1}{\log_2 M}\right) \\
&\quad \times \left(\sup_{\substack{P_1, P_2, \ldots, P_k \ge 0, \\ \sum_{i=1}^{k} P_i > 0}} \min_{j \in \mathcal{R}} \frac{\sup_{\substack{P_{\mathbf{x}}: \\ \mathbb{E}[|x_i|^2] \le P_i}} I(\mathbf{x}; y_j | \mathbf{h}_j)}{\sum_{i=1}^{k} P_i} \right)^{-1} && (127)
\end{aligned}$$

where to get (127), we have substituted E_i/n by P_i in (126) and taken supremum over P_i for all $i = 1, \ldots, k$.

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

$$\lim_{M \to \infty} \sup \frac{E_{\text{total}}}{\log_2 M} < E_b + \delta. \quad (128)$$

This implies that

$$\begin{aligned}
E_b \ge& \lim_{M \to \infty} \sup \left(1 - \epsilon - \frac{1}{\log_2 M}\right) \\
&\times \left(\sup_{\substack{P_1, \ldots, P_k \ge 0, \\ \sum_{i=1}^{k} P_i > 0}} \min_{j \in \mathcal{R}} \frac{\sup_{\substack{P_{\mathbf{x}}: \\ \mathbb{E}[|x_i|^2] \le P_i}} I(\mathbf{x}; y_j | \mathbf{h}_j)}{\sum_{i=1}^{k} P_i} \right)^{-1}. \quad (129)
\end{aligned}$$

Moreover, if E_b is an achievable energy per bit value, then we can take supremum of the left-hand side of (129) over all $0 < \epsilon < 1$ to get

$$E_b \ge \inf_{\substack{P_1, \ldots, P_k \ge 0, \\ \sum_{i=1}^{k} P_i > 0}} \max_{j \in \mathcal{R}} \frac{\sum_{i=1}^{k} P_i}{\sup_{\substack{P_{\mathbf{x}}: \\ E[|x_i|^2] \le P_i}} I(\mathbf{x}; y_j | \mathbf{h}_j)}. \quad (130)$$

Therefore, $E_{b\min}$ should also satisfy (130) implying (9). □

APPENDIX II
PROOF OF LEMMA 2

1) Proof: First, note that

$$P[\mathcal{D}_k^c] = P\Bigg[\exists C \in \mathcal{C}: \nu(C) \notin \left[(1-\delta)\frac{(k-1)s^2}{A_k}, (1+\delta)\frac{(k-1)s^2}{A_k}+1\right)\Bigg] \le P\left[\exists C \in \mathcal{C}: \nu(C) < (1-\delta)\frac{(k-1)s^2}{A_k}\right] + P\left[\exists C \in \mathcal{C}: \nu(C) \ge (1+\delta)\frac{(k-1)s^2}{A_k}+1\right] \quad (131)$$

by taking union bound over the two ways of violating the condition for $\mathcal{D}_k$. Using union bond again, each of the two terms in the right-hand side of (131) can be further bounded by the sum of probabilities of violations for all cells. We note that all the cells are identical as far as number of nodes in them is concerned, except for the fact that the origin cell, say C_1, already contains an additional node (the source node). Therefore, the origin cell will have the maximum probability of having greater than $((1+\delta)(k-1)s^2/A_k)+1$ nodes and the nonorigin cells will have the maximum probability of having less than $(1-\delta)(k-1)s^2/A_k$ nodes. Let C_2 be a representative cell for nonorigin cells. Noting that there are A_k/s^2 cells, we have the following union bound on (131):

$$P[\mathcal{D}_k^c] \le \frac{A_k}{s^2}\left(P\left[\nu(C_1) \ge (1+\delta)\frac{(k-1)s^2}{A_k}+1\right] + P\left[\nu(C_2) < (1-\delta)\frac{(k-1)s^2}{A_k}\right]\right). \quad (132)$$

To further bound the probabilities in (132), we use the Chernoff bound. Let X_i be the indicator function of node i falling in C_2. Then, for $i = 2, \ldots, k$

$$X_i = \begin{cases} 0, & \text{w.p. } 1-\dfrac{s^2}{A_k} \\ 1, & \text{w.p. } \dfrac{s^2}{A_k} \end{cases}$$

are independent random variables. The probability that cell C_2 contains fewer than $(1-\delta)(k-1)s^2/A_k$ nodes is equivalent to evaluating

$$P\left[\sum_{i=2}^{k} X_i < (1-\delta)\frac{(k-1)s^2}{A_k}\right]. \quad (133)$$

By a simple change of variables $X_i' = X_i - (s^2/A_k)$, we get that

$$X_i' = \begin{cases} -\dfrac{s^2}{A_k}, & \text{w.p. } 1-\dfrac{s^2}{A_k} \\ 1-\dfrac{s^2}{A_k} & \text{w.p. } \dfrac{s^2}{A_k} \end{cases}, \qquad i = 2, \ldots, k$$

which satisfies [2, Assumption A.1.3]. So, by the Chernoff bound [2, Th. A.1.13]

$$P\left[\sum_{i=2}^{k} X_i' < (-\delta)\frac{(k-1)s^2}{A_k}\right] < \exp\left(-\delta^2\frac{(k-1)s^2}{2A_k}\right). \quad (134)$$

The right-hand side of (134) provides an upper bound on (133) and thus an upper bound on the probability of $\nu(C_2)$ being less than $(1-\delta)(k-1)s^2/A_k$.

Retaining the variable X' and applying [2, Th. A.1.11], we also get

$$P\left[\nu(C_1) \ge (1+\delta)\frac{(k-1)s^2}{A_k}+1\right] = P\left[\sum_{i=2}^{k} X_i' \ge \delta\,\frac{(k-1)s^2}{A_k}\right] < \exp\left(-\delta^2(1-\delta)\frac{(k-1)s^2}{2A_k}\right). \quad (135)$$

Using (134) and (135) in (132) and noticing that the right-hand side of (135) is greater than the right-hand side of (134), we get the final result. □

APPENDIX III
PROOF OF LEMMA 3

Proof: Label the k cells as $C_1, C_2, \ldots, C_k$, where C_1 is the origin cell. Let X_i be a random variable indicating the cell in which the nonsource node i falls. We will also use X_i to indicate whether the node i falls in the window portion or the nonwindow portion of the cell, in the following manner. Let X_i take the integer values from $-k$ to $+k$ excluding zero. If $X_i = +j$ for some $j \in \{1, \ldots, k\}$, it implies that the node i falls into the windowed portion (with the area β^2/λ) of C_j. If $X_i = -j$ for some $j \in \{1, \ldots, k\}$, it implies that the node i falls into the nonwindowed portion (with the area $(1-\beta^2)/\lambda$) of C_j. Clearly, the X_is are i.i.d. with common distribution X given by

$$X = \begin{cases} +j, & \text{w.p. } \dfrac{\beta^2}{k} \\ -j, & \text{w.p. } \dfrac{1-\beta^2}{k} \end{cases}, \qquad \text{for all } j \in \{1, \ldots, k\}. \quad (136)$$

Though all the other cells are identical, C_1 already contains the source node outside its window, so for the next set of calculations, we will only be dealing with the $k-1$ nonsource nodes and $k-1$ nonorigin cells.

Consider the function $f: \{-k, \ldots, -1, +1, \ldots, k\}^{k-1} \mapsto \mathbb{N}$, that counts the number of good nonorigin cells in a realization of $(X_2, X_3, \ldots, X_k)$, i.e.,

$$f(x_2, x_3, \ldots, x_k) = \sum_{i=2}^{k} \mathbf{1}\Big\{\exists j \in \{2, \ldots, k\} : x_j = +i, \text{ and } x_\ell \ne \pm i \,\forall \ell \in \{2, \ldots, j-1, j+1, \ldots, k\}\Big\}. \quad (137)$$

Taking expectation of the function f, we get

$$\begin{aligned}&\mathbb{E}[f(X_2,\dots,X_k)]\\&=\sum_{i=2}^{k}\mathbb{E}\Big[\mathbf{1}\Big\{\exists j\in\{2,\dots,k\}:X_j=+i,\\&\qquad\text{and } X_\ell\neq\pm i\,\forall\ell\in\{2,\dots,j-1,j+1,\dots,k\}\Big\}\Big]\\&=(k-1)\,P\Big[\exists j\in\{2,\dots,k\}:X_j=+2,\\&\qquad\text{and } X_\ell\neq\pm 2\,\forall\ell\in\{2,\dots,j-1,j+1,\dots,k\}\Big]\end{aligned} \tag{138}$$

where (138) follows by observing that all the nonorigin cells are identical; so, in (138), we pick cell 2 as a representative nonorigin cell. Using (136), we can evaluate the probability in (138) which is the probability that X_j is $+2$ for exactly one $j\in\{2,\dots,k\}$ (which can happen in $k-1$ ways) and the rest $k-2$ nodes fall outside the cell 2. Therefore

$$\mathbb{E}[f(X_2,\dots,X_k)]=\beta^2\frac{(k-1)^k}{k^{k-1}}. \tag{139}$$

Our task is to bound the deviation of the function f from its mean value. We do this using McDiarmid's inequality, which is a generalization of the Chernoff bound:

Lemma 4 (McDiarmid's Inequality [16]): Let $X_1,X_2,\dots,X_n$ be independent random variables, with X_i taking values in a set $\mathcal{X}$ for each i. Suppose that the function $f:\mathcal{X}^n\mapsto\mathbb{R}$ satisfies

$$|f(\mathbf{x})-f(\mathbf{x}')|\le c_i \tag{140}$$

whenever $\mathbf{x}$ and $\mathbf{x}'$ differ only in the i^{th} coordinate. Then, for any $t>0$

$$\begin{aligned}&P\left[|f(X_1,X_2,\dots,X_n)-\mathbb{E}[f(X_1,X_2,\dots,X_n)]|\ge t\right]\\&\qquad\le 2\exp\left(-\frac{2t^2}{\sum_{i=1}^{n}c_i^2}\right).\end{aligned} \tag{141}$$

In our case, the value of function f varies by at most 2 whenever only one component changes. This maximal variation corresponds to those situations when a node belonging to a good cell moves to another good cell, thus destroying the "goodness" of both of them. Similar situation can be imagined when two good cells are generated by relocation of one node. Therefore

$$c_i\le 2,\qquad \text{for all } i=2,3,\dots,k. \tag{142}$$

Setting $t=\delta\beta^2(k-1)^k/k^{k-1}$ and using the value of $\mathbb{E}[f(X_2,\dots,X_k)]$ from (139) in Lemma 4, we immediately get (76). ☐

Appendix IV
Proof of Theorem 4

1) Proof: The analysis is divided into three separate cases.
1) $r_0<(1-\beta)s$.
2) $(k-1)s^2<r_0^2$.
3) $(1-\beta)s\le r_0\le\sqrt{k-1}s$

In each of the cases above, we derive a lower bound on $\frac{E_b}{N_0}_{\min}$, and also propose a flooding algorithm with suitable parameters to get an upper bound on $\frac{E_b}{N_0}_{\text{flood}}$.

Case 1: Let us first upper bound $\sum_{j\in\mathcal{R}\setminus\{i\}}g(r_{ij})$ for any node i and the destination set $\mathcal{R}=\{2,\dots,k\}$. Begin by noticing that the number of nodes within ℓ steps (horizontal, vertical, or diagonal) of any node is at most 8ℓ. Moreover, the distance to any node in a cell ℓ steps away is at least $(\ell-1)s+(1-\beta)s\ge\ell(1-\beta)s$ and there are at most $\sqrt{k}-1$ steps in any direction. Therefore, for node i

$$\begin{aligned}\sum_{j\in\mathcal{R}\setminus\{i\}}g(r_{ij})&\le\sum_{\ell=1}^{\sqrt{k}-1}8\ell\,g(\ell(1-\beta)s)\\&=8\sum_{\ell=1}^{\infty}\frac{\ell}{\ell^\alpha(1-\beta)^\alpha s^\alpha}\\&\le\frac{8\zeta(\alpha-1)}{(1-\beta)^\alpha s^\alpha}.\end{aligned} \tag{143}$$

Hence, $G(\mathcal{R})$ satisfies

$$G(\mathcal{R})\le\frac{8\zeta(\alpha-1)}{(1-\beta)^\alpha s^\alpha(k-1)}. \tag{144}$$

Therefore

$$\frac{E_b}{N_0}_{\min}\ge\frac{(1-\beta)^\alpha s^\alpha(k-1)\log_e 2}{8\zeta(\alpha-1)} \tag{145}$$

by Theorem 1.

Next, we note that any node has a sequence (of length at most $\sqrt{k}$) of adjacent (horizontal, vertical, or diagonal) cells that begins at the origin cell and ends at the cell containing the node. This translates into a sequence of nodes such that any two adjacent nodes are within a distance $\sqrt{8}s$ of each other. Thus, the multihop scheme

$$\texttt{FLOOD}\left(\frac{N_0\log_e 2}{g(\sqrt{8}s)}+\epsilon_1,\frac{N_0\log_e 2}{g(\sqrt{8}s)}+\epsilon_1\right) \tag{146}$$

for any $\epsilon_1>0$, would work well. The energy consumption of this scheme is at most $\left(kN_0\log_e 2/g(\sqrt{8}s)\right)+k\epsilon_1$. Since $\sqrt{8}s>r_0$ by the defining condition for this case, we have $g(\sqrt{8}s)=8^{-\alpha/2}s^{-\alpha}$.

Therefore, the energy consumption per bit of (146) is bounded as

$$E_{b\text{total}}\le\frac{2^{\frac{3\alpha}{2}+4}\zeta(\alpha-1)}{(1-\beta)^\alpha}E_{b\min}+\epsilon \tag{147}$$

for any $\epsilon>0$. Therefore, $E_{b\text{flood}}$ which is the infimum of all the achievable $E_{b\text{total}}$ values satisfies

$$E_{b\text{flood}}\le\frac{2^{\frac{3\alpha}{2}+4}\zeta(\alpha-1)}{(1-\beta)^\alpha}E_{b\min}. \tag{148}$$

Case 2: If $(k-1)s^2<r_0^2$, then $ks^2<2r_0^2$ for any $k>1$. This implies that the network can be contained within a square of side $\sqrt{2}r_0$. Therefore, the maximum distance between any two nodes is at most $2r_0$. So, any node can be reached by the one shot transmission scheme

$$\texttt{FLOOD}\left(\frac{N_0\log_e 2}{g(2r_0)}+\epsilon_1, 0\right) \tag{149}$$

for any $\epsilon_1 > 0$. Note that $g(2r_0) = 2^{-\alpha}r_0^{-\alpha}$.

For the lower bound on energy per bit, note that the gain to any node cannot exceed $\bar{g}$ and there are $k-1$ destination nodes. Therefore, the effective network loss is greater than $1/\bar{g}$. Hence

$$\frac{E_b}{N_0}_{\min} \geq \frac{\log_e 2}{\bar{g}} \tag{150}$$

by Theorem 1. This immediately implies that

$$E_{b\text{total}} \leq 2^\alpha r_0^\alpha \bar{g}\, E_{b\min} + \epsilon \tag{151}$$

for any $\epsilon > 0$. Therefore

$$E_{b\text{flood}} \leq 2^\alpha r_0^\alpha \bar{g}\, E_{b\min}. \tag{152}$$

Case 3: Define

$$L \triangleq \left\lfloor \frac{r_0}{(1-\beta)s}\right\rfloor. \tag{153}$$

Since $L \geq 1$, we have the following bounds on L:

$$\frac{1}{2}\frac{r_0}{(1-\beta)s} < L \leq \frac{r_0}{(1-\beta)s} \tag{154}$$

and

$$\frac{r_0}{(1-\beta)s} < L+1 \leq \frac{2r_0}{(1-\beta)s}. \tag{155}$$

Next, by the same argument as in Case 1

$$\sum_{j\in\mathcal{R}\setminus\{i\}} g(r_{ij}) \leq \sum_{\ell=1}^{\sqrt{k}-1} 8\ell\, g(\ell(1-\beta)s). \tag{156}$$

Continuing with (156), we get the following steps [the explanation of these steps is similar to that of (37)–(40)]:

$$\begin{aligned}\sum_{j\in\mathcal{R}\setminus\{i\}} g(r_{ij}) &\leq 8\sum_{\ell=1}^{L}\ell g\left(\ell(1-\beta)s\right) + 8\sum_{\ell=L+1}^{\infty}\ell g\left(\ell(1-\beta)s\right)\\ &\leq 4\bar{g}L(L+1) + 8\sum_{\ell=L+1}^{\infty}\frac{\ell}{\ell^\alpha(1-\beta)^\alpha s^\alpha}\\ &\leq 4\bar{g}\frac{2r_0^2}{(1-\beta)^2s^2} + \frac{8}{(1-\beta)^\alpha(\alpha-2)}\frac{1}{s^\alpha}\frac{1}{L^{\alpha-2}}\\ &\leq 8\bar{g}\frac{r_0^2}{(1-\beta)^2s^2}\\ &\quad + \frac{8}{(1-\beta)^\alpha(\alpha-2)}\frac{1}{s^\alpha}\frac{(2(1-\beta))^{\alpha-2}s^{\alpha-2}}{{r_0}^{\alpha-2}}\\ &= \frac{8}{(1-\beta)^2s^2}\left(\bar{g}r_0^2 + \frac{2^{\alpha-2}}{(\alpha-2)r_0^{\alpha-2}}\right)\\ &= \frac{c_2}{s^2}\end{aligned} \tag{157}$$

where we set

$$c_2 = \frac{8}{(1-\beta)^2}\left(\bar{g}r_0^2 + \frac{2^{\alpha-2}}{(\alpha-2)r_0^{\alpha-2}}\right). \tag{158}$$

By (157), the effective network loss is at least $(k-1)s^2/c_2$. Therefore, by Theorem 1

$$\frac{E_b}{N_0}_{\min} \geq \frac{(k-1)s^2\log_e 2}{c_2}. \tag{159}$$

For the achievable part, consider the algorithm

$$\texttt{FLOOD}\left(\frac{N_0\log_e 2}{g(2\sqrt{2}sL)}+\epsilon_1, \frac{N_0\log_e 2}{\sum_{\ell=1}^{L}\ell\, g(2\sqrt{2}s\ell)}+\epsilon_1\right) \tag{160}$$

for any $\epsilon_1 > 0$.

Before analyzing its energy consumption, let us first see why should this algorithm work. The maximum distance between two nodes belonging to cells that are $\ell \geq 1$ vertical, horizontal, or diagonal steps away is $(\ell+1)\sqrt{2}s \leq 2\sqrt{2}s\ell$. Therefore, transmitting with energy per bit $\left(N_0\log_e 2/g(2\sqrt{2}sL)\right)+\epsilon_1$ ensures that any node within L steps can decode the message reliably. This is what the source node does. Hence, by the end of the first time slot, any cell belonging to the set

$$\mathcal{S}_1 \triangleq \{C(x,y) : \max\{x,y\} \leq (L-1)+1\} \tag{161}$$

can decode the message reliably. For any $T < T_k \triangleq \sqrt{k}-L+1$, define

$$\mathcal{S}_T \triangleq \{C(x,y) : \max\{x,y\} \leq (L-1)+T\}. \tag{162}$$

Suppose that by the end of time slot $T < T_k - 1$, the set of cells which have decoded the message is a superset of $\mathcal{S}_T$. We claim that for any node in the set of cells $\mathcal{S}_{T+1}\setminus\mathcal{S}_T$, the total received energy per bit by the end of time slot $T+1$ due to transmissions from nodes in $\mathcal{S}_T$ is greater than $N_0\log_e 2$. This is true since for any $\ell \leq L$, any node in the set of cells $\mathcal{S}_{T+1}\setminus\mathcal{S}_T$ has at least ℓ distinct nodes in $\mathcal{S}_T$ which are exactly ℓ vertical, horizontal, or diagonal steps away. Note that these nodes at a step distance of ℓ are at most $2\sqrt{2}s\ell$ distance away. Since all the nodes in $\mathcal{S}_T$ have transmitted by the end of time slot $T+1$ with an energy per bit of $\left(N_0\log_e 2/\sum_{\ell=1}^{L}\ell\, g(2\sqrt{2}s\ell)\right)+\epsilon_1$, the total received energy per bit at any node in $\mathcal{S}_{T+1}\setminus\mathcal{S}_T$ is greater than $N_0\log_e 2$. Thus, the nodes in $\mathcal{S}_{T+1}$ are covered and by induction, all the nodes are covered by the end of time slot T_k.

Since $L \leq r_0/((1-\beta)s)$, for all $\ell \leq L$

$$\begin{aligned}g(2\sqrt{2}s\ell) &\geq g(2\sqrt{2}sL)\\ &\geq \left(2\sqrt{2}\frac{r_0}{(1-\beta)}\right)^{-\alpha}\\ &= \frac{(1-\beta)^\alpha}{8^{\alpha/2}r_0^\alpha}.\end{aligned} \tag{163}$$

Using (163), the total energy consumption of (160) can be bounded by

$$\begin{aligned} E_{b\text{total}} &\leq \frac{N_0 \log_e 2}{g(2\sqrt{2}sL)} + \frac{(k-1)N_0 \log_e 2}{\sum_{\ell=1}^{L} \ell\, g(2\sqrt{2}s\ell)} + k\epsilon_1 \\ &\leq \frac{8^{\alpha/2} r_0^{\alpha} N_0 \log_e 2}{(1-\beta)^{\alpha}} \left(1 + \frac{2(k-1)}{L(L+1)}\right) + k\epsilon_1 \\ &\leq \frac{8^{\alpha/2} r_0^{\alpha} N_0 \log_e 2}{(1-\beta)^{\alpha}} \left(1 + \frac{4(1-\beta)^2 s^2 (k-1)}{r_0^2}\right) + k\epsilon_1 \end{aligned} \tag{164}$$

for any $\epsilon_1 > 0$. As before, taking the infimum of all the $E_{b\text{total}}$ values removes the $+k\epsilon_1$ term from the right-hand side of (164) to yield a bound on $E_{b\text{flood}}$.

Therefore, from (159) and (164), the upper and lower bounds are related by

$$\begin{aligned} E_{b\text{flood}} &\leq \frac{8^{\alpha/2} r_0^{\alpha} c_2}{(1-\beta)^{\alpha}} \left(\frac{1}{s^2(k-1)} + \frac{4(1-\beta)^2}{r_0^2}\right) E_{b\min} \\ &\leq \frac{8^{\alpha/2} r_0^{\alpha} c_2}{(1-\beta)^{\alpha}} \left(\frac{1}{r_0^2} + \frac{4(1-\beta)^2}{r_0^2}\right) E_{b\min} \qquad (165) \\ &= \frac{8^{\frac{\alpha}{2}+1} r_0^{\alpha}}{(1-\beta)^{\alpha+2}} \left(\bar{g} + \frac{2^{\alpha-2}}{(\alpha-2) r_0^{\alpha}}\right) \left(1 + 4(1-\beta)^2\right) \times \\ &\quad E_{b\min}. \qquad (166) \end{aligned}$$

Note that we have used the condition $(k-1)s^2 \geq r_0^2$ in (165).

Finally, putting together the results of all three cases, we get the statement of Theorem 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] N. Alon and J. Spencer, *The Probabilistic Method*, 2nd ed. New York: Wiley, 2000.

[3] X. Cai, Y. Yao, and G. Giannakis, "Achievable rates in low-power relay links over fading channels," *IEEE Trans. Commun.*, vol. 53, no. 1, pp. 184–194, Jan. 2005.

[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. Cover and A. E. Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. IT-25, no. 5, pp. 572–584, Sep. 1979.

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

[7] A. F. Dana and B. Hassibi, "On the power efficiency of sensory and ad hoc wireless networks," *IEEE Trans. Inf. Theory*, vol. 52, no. 7, pp. 2890–2914, Jul. 2006.

[8] A. El Gamal and J. Mammen, "Optimal hopping in ad hoc wireless networks," in *Proc. 25th Annu. Joint Conf. IEEE Comput. Commun. Soc.*, Barcelona, Spain, Apr. 2006, pp. 1–10.

[9] A. El 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] A. K. Haddad, V. Ribeiro, and R. Riedi, "Broadcast capacity in multihop wireless networks," in *Proc. 12th Annu. Int. Conf. Mobile Comput. Netw.*, Los Angeles, CA, Sep. 2006, pp. 239–250.

[11] 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.

[12] J. Laneman, D. Tse, and G. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.

[13] 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.

[14] H. Lim and C. Kim, "Multicast tree construction and flooding in wireless ad hoc networks," in *Proc. 3rd ACM Int. Workshop Model. Anal. Simul. Wireless Mobile Syst.*, Boston, MA, Aug. 2000, pp. 61–68.

[15] 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.

[16] C. McDiarmid, "On the method of bounded differences," *Surv. Combin.*, vol. 141, pp. 148–188, 1989.

[17] B. S. Mergen and M. Gastpar, "The scaling of the broadcast capacity of extended wireless networks with cooperative relays," in *Proc. 41st Asilomar Conf. Signals Syst. Comput.*, Pacific Grove, CA, Nov. 2007, pp. 873–877.

[18] B. Sirkeci-Mergen and M. Gastpar, "On the broadcast capacity of wireless networks with cooperative relays," *IEEE Trans. Inf. Theory*, vol. 56, no. 8, pp. 3847–3861, Aug. 2010.

[19] 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.

[20] B. S. Mergen, A. Scaglione, and G. Mergen, "Asymptotic analysis of multistage cooperative broadcast in wireless networks," *IEEE Trans. Inf. Theory*, vol. 52, no. 6, pp. 2531–2550, Jun. 2006.

[21] K. Obraczka, K. Viswanath, and G. Tsudik, "Flooding for reliable multicast in multi-hop ad hoc networks," *Wireless Netw.*, vol. 7, no. 6, pp. 627–634, Nov. 2001.

[22] A. Scaglione and Y. Hong, "Opportunistic large arrays: Cooperative transmission in wireless multihop ad hoc networks to reach far distances," *IEEE Trans. Signal Process.*, vol. 51, no. 8, pp. 2082–2092, Aug. 2003.

[23] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity—Part I: System description," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1927–1938, Nov. 2003.

[24] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity—Part II: Implementation aspects and performance analysis," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1939–1948, Nov. 2003.

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

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

[27] 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, Mar. 2000, vol. 2, pp. 585–594.

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

[29] R. Zheng, "Information dissemination in power-constrained wireless networks," in *Proc. 25th Annu. Joint Conf. IEEE Comput. Commun. Soc.*, Barcelona, Spain, Apr. 2006, pp. 1–10.

Aman Jain received the B.Tech. degree in electrical engineering from the Indian Institute of Technology, Kanpur, India, in 2005 and the M.A. degree in electrical engineering from Princeton University, Princeton, NJ, in 2007, where he is currently working towards the Ph.D. degree.

In 2004, he was a summer intern at the Embedded Systems Laboratory, University of Leicester, Leicester, U.K. In 2009, he was a Graduate Fellow with the McGraw Center for Teaching and Learning, Princeton University. His research interests include information theory, wireless communications, networks, and signal processing.

Mr. 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. 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 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.

Dr. 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 the 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 the Editor-in-Chief of the *Foundations and Trends in Communications and Information Theory*.