Overcoming Untuned Radios in Wireless Networks With Network Coding

Dragan Petrović, *Student Member, IEEE*, Kannan Ramchandran, *Fellow, IEEE*, and Jan Rabaey, *Fellow, IEEE*

***Abstract*—The drive toward the implementation and massive deployment of wireless sensor networks calls for ultralow-cost and low-power nodes. While the digital subsystems of the nodes are still following Moore's Law, there is no such trend regarding the performance of analog components. This work proposes a fully integrated architecture of both digital and analog components (including local oscillator) that offers significant reduction in cost, size, and overall power consumption of the node. Even though such a radical architecture cannot offer the reliable tuning of standard designs, it is shown that by using random network coding, a dense network of such nodes can achieve throughput linear in the number of channels available for communication. Moreover, the ratio of the achievable throughput of the untuned network to the throughput of a tuned network with perfect coordination is shown to be close to $1/e$. This work uses network coding to leverage the fact that throughput equal to the max-flow in a graph is achievable even if the topology is not know *a priori*. However, the challenge here is finding the max-flow of the random graph corresponding to the network.**

***Index Terms*—Network coding, sensor networks, untuned radios.**

I. Introduction

THE emerging field of wireless sensor networks has become a very active area of both academic research [1]–[4] and industrial development [5]–[9]. The potential scenarios for use of sensor networks are far ranging. Some of the near-term applications include monitoring the structural integrity of buildings and bridges [10], environmental control within living and working spaces [11], habitat monitoring in animal sanctuaries [12], highway traffic control [13], and warehouse inventory tracking [14]. The potential applications that are being considered in the long-term include such science-fiction-like concepts as smart surfaces that can respond to contact or serve as a communication backplane, as well as airplane wings that can provide real-time monitoring of and alerts regarding the state of every square centimeter of the wing surface. It may even be possible to eventually produce sensor nodes small enough to be inserted into the blood-stream to provide real-time diagnostics of factors such as blood pressure, blood flow, glucose and insulin levels, etc.

To make such deployments economically and technologically feasible, it is necessary to drastically reduce the cost, size and energy consumption of the nodes available today. Moore's Law still provides for exponential reduction of these metrics over time when it comes to the digital components that comprise the memory, computation and coding of the nodes. However, there is no equivalent trend to Moore's Law that applies to the analog components needed for the radios that enable the nodes to communicate with one another. This work introduces an architecture for the analog radios that is expected to reduce the cost, size and energy consumption of the nodes by an order of magnitude. In fact, the proposed architecture can allow the energy consumption of the nodes to be so low that they could be fully powered by energy scavenged from the environment [15]. The penalty for using such a radical architecture is that the radios become untuned and it is no longer possible to guarantee that any arbitrary pair of nodes will be able to communicate with each other. Instead, it becomes necessary to rely on the density of nodes to make the overall network capable of providing reliable communication.

Narrowband radios have proven to be the architecture of choice for low-power applications [6], [7], [16], as they are low in complexity and consume less power than spread spectrum or other wide-band techniques. One fundamental requirement of narrowband radios is that the transmitter's carrier frequency and the receiver's detection frequency must be well-matched. This is traditionally accomplished by employing a crystal at both the transmitter and receiver to provide the same low frequency reference. This reference frequency is multiplied via a phase-locked loop (PLL) to generate the carrier wave. However, the off-chip crystal contributes significantly to the cost, size, and power consumption of such transceivers. The cost is due to the external quartz crystals being more expensive than the silicon used for the baseband signal processing as well as the need to bond separate components. This problem is especially acute in the design of highly integrated transceivers for wireless sensor networks. The size of traditional low power transceivers is largely due to the external crystal reference and the interface between the crystal and the silicon integrated circuit (IC). Additionally, the power consumption of low power radios is dominated by the crystal referenced PLL. Therefore, great savings in all three of these areas could be obtained by eliminating the off-chip crystal and PLL.

Even when care is taken to ensure that all radios are tuned and are attempting to communicate on the same frequency, reliable communication is not guaranteed. Practical implementations of sensor networks are notorious for having unstable links because narrowband communication is susceptible to deep fades between nodes [17], [18]. Since it is not feasible to overcome these fades by transmitting with more power (due

Manuscript received March 15, 2005; revised February 14, 2006. This work was sponsored in part by NSF grants CCR-0325311 and CCR-0330514.

The authors are with the Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, CA 94704 USA (e-mail: dragan@eecs.Berkeley.edu; kannanr@eecs.Berkeley.edu; jan@eecs.Berkeley.edu).

Communicated by R. Koetter, Guest Editor.
Digital Object Identifier 10.1109/TIT.2006.874396

to power-constraints), it has been proposed that randomized algorithms be used to ensure reliable communication [19], [20]. These randomized algorithms exploit the broadcast nature of wireless transmissions to provide reliable multi-hop communication. The key idea is for a transmitting node to send a beacon to many potential forwarding nodes and then select one node to be the next hop for the packet among those that respond to the beacon. However, collisions among the responses to the beacon as well as the time-varying quality of the communication channels (a channel may be good during the beaconing, but become bad during the response and/or data transmission) contribute significant overhead to such schemes.

This paper proposes a fundamentally different way of designing and operating a transceiver. The quartz crystal is eliminated and replaced by an on-chip resonator such as an inductor–capacitor (*LC*)-circuit or a nanoelectromechanical resonant structure. This makes it possible to economically produce millions of nodes and densely deploy them by weaving them into fabrics or mixing them with paint. The proposed architecture allows a sensor node to be developed entirely out of thin-film technologies (radio, digital, battery, energy scavenging, and sensing). However, the drawback of such architectures is that the variations in the manufacturing process are large, resulting in untuned radios. Therefore, two narrowband radios produced by such a process are not likely to be able to communicate with each other. To address this problem, a low-complexity communication protocol is proposed that makes use of the high density of nodes to ensure reliable communication using such untuned radios even without the need for handshaking protocols or re-transmission. By eliminating the need for this kind of coordination, the protocol is also made more robust to link failures, while the density that is made possible by such low cost designs makes the network robust to the failure of individual nodes.

A. Main Results

We consider using the nodes to form a communication backplane carrying data between a source and a destination. The data is transported in a multihop fashion by a network of nodes that employ untuned narrowband radios. Let N denote the number of unit-capacity channels available for communication. We show that by using random linear network coding, achievable throughput of the network is $\Theta(N)$, same as in a fully-coordinated network of tuned radios. Moreover, the ratio of the achievable throughput of the untuned network to the achievable throughput of a tuned network is shown to be close to $1/e$. In contrast, simple random routing (in which forwarding nodes are only allowed to randomly select one of the packets to forward, rather than combining the packets they receive to form the output packet) has constant throughput, even as N grows. In other words, when employing simple random routing, the throughput per channel goes to zero as the number of available channels grows.

This work uses network coding to leverage the fact that throughput equal to the max-flow in a graph is achievable even when the connectivity of the network is not known *a priori*. However, the challenge here is in finding the max-flow of the random graph corresponding to the network.

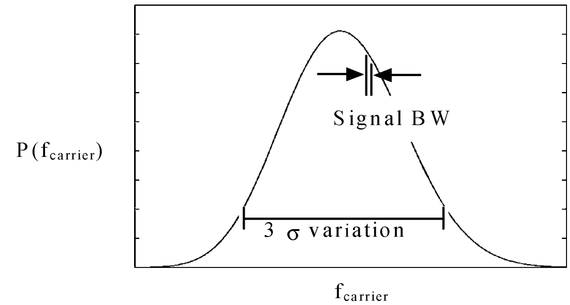

Fig. 1. Signal bandwidth relative to process variation when using on-chip *LC* resonators to provide the carrier frequency for narrowband radios. When using an *LC* resonator without a crystal reference, it is not possible for a radio to know exactly at what frequency it is operating.

II. Untuned Radios

The drawback of using an on-chip resonator is that the variations in the manufacturing process are large meaning that the distribution of the resonant frequency of the manufactured oscillators will have large variance.

Fig. 1 presents a qualitative illustration of the challenges presented by this approach. In a traditional narrow-band architecture with a quartz crystal reference, the signal bandwidth is typically orders of magnitude larger than the center frequency tolerance. When using untuned receivers, the situation is reversed and the carrier frequency variation is orders of magnitude greater than the signal bandwidth [21], [22]. To achieve frequency tolerance approaching a quartz crystal, prohibitively expensive trimming would have to take place. In addition, drift over time, temperature, and supply voltage would quickly render the trimming inaccurate. The other option is to leave the transmitters and receivers untuned, which means that two narrowband radios produced by such a process are not likely to be able to communicate with each other. If the input frequency range of a particular receiver did admit the transmitter's carrier frequency, the transmission would be received successfully. Otherwise, the result would be the same as if the transmitter were communicating on a channel orthogonal to the one that the receiver is monitoring. This means that the unreliable manufacturing process results in having multiple channels available for communication but neither the transmitting nor receiving nodes can select on which particular channel to communicate. Instead, the channel on which a transmitter or receiver communicate are random.

Even though a particular transmit–receive pair may not be able to communicate because they would be effectively tuned to different channels, a sufficiently high density of nodes ensures a high probability that there are pairs of transmitters and receivers that can communicate with one another. The number of channels available in the system is determined by the ratio of the variations of the manufacturing process to the receiver bandwidth. As long as the bandwidth admitted by the receive filters is greater than the signal bandwidth, there will be a nonzero probability that two randomly selected nodes will be able to communicate with each other.

In the rest of the paper, the abstraction of having multiple channels available for communication is made. Note that it is not necessary for the channels to be orthogonal (in fact, they are not!). What is important is the number of transmitter carrier frequencies that fall within the bandwidth being monitored by a particular receiver. The probability that any given transmitter falls within the receiver's range is dependent only on the ratio of the range of possible carrier frequencies to the receiver bandwidth. This ratio is equivalent to the number of channels in our analysis because it is equal to the maximum number of independent transmissions that can be made simultaneously without interfering with one another.

To maximize the throughput of the network, it is necessary to maximize the probability that a channel is occupied by exactly one transmitter during communication. This will maximize the number of channels that contain a decodable transmission. It can be shown that when there are N channels available for communication and each transmitter is independently and randomly assigned to a channel with the same probability distribution, the probability that a channel is occupied by exactly one transmitter is maximized when there are exactly N transmitters and each transmitter is equally likely to be assigned to any of the channels. In this case, the probability that a channel contains exactly one transmission is asymptotically, for large N, equal to $1/e$ (the result is proven in Appendix A). This implies that, in order to maximize the throughput, the network should be operated with N active transmitters within communication range of each other, in which case each transmission will experience a collision with probability $1 - 1/e$.

Having N active transmitters within communication range maximizes the probability that a receiver will have exactly one transmitter in the range of frequencies it monitors; however, it may still be possible that a transmitter occupies a unique channel, but no receiver is tuned to that channel. In order to increase the probability that a noncolliding transmission is heard as well as the probability that a receiving node hears at least one packet, each one is equipped with several receive radios, with each radio tuned to a different channel (by using a different *LC*-circuit as the local oscillator for each radio – this also allows the nodes to transmit on different random channels at different times by selecting any of its *LC*-circuits to provide the carrier frequency for the transmitter).

Denote the number of receive-radios on each node with L. In Section IV, we derive the relationship between the value of L and the throughput of the network. We show that it is possible to achieve throughput that is linear in the number of channels even with a constant value of L. It is important to show that this is achievable with a constant L because requiring L to grow with the number of channels would correspond to requiring more hardware on each node, and this is exactly what we are trying to avoid. In order to give guidelines for practical deployments, simulations are used to estimate the throughput that can be achieved with practical values of L.

III. Multihop Communication

We consider using the nodes to form a communication backplane carrying data between a source and a destination. This scenario is shown in Fig. 2. The source sends its data to the destination in a multihop fashion using geographic routing. The region between the source and the destination is divided into blocks such that any two nodes in adjacent blocks are within communication range of each other. This allows the nodes to make use of the broadcast nature of wireless communication by having multiple potential forwarders for each transmission. In previous works, the broadcast nature of the wireless channel is exploited by using beaconing to establish a connection between a transmitting node and a forwarder [19], [20], but in our scheme, the transmitters simply broadcast their data and assume that at least one node in the next block will receive and relay the packet.

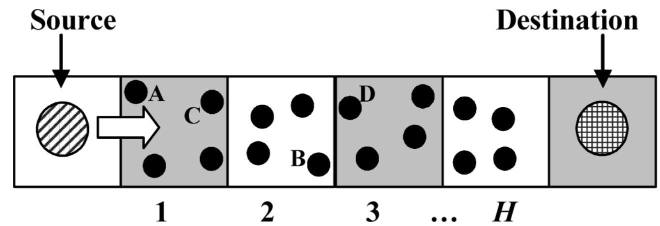

Fig. 2. Proposed multi-hop communication method. Nodes in each block listen to L randomly selected frequency bands. If a node detects transmissions from the previous block in any of those bands, it combines the inputs using random linear network coding and broadcasts the result to the next block.

In the proposed scheme, anytime a node transmits a packet, it includes in the packet's header the coordinates of the next block toward the destination as well as the coordinates of the destination so that only nodes in the next block will relay the information, again specifying the direction in which the data should propagate. It is assumed that the nodes know their position [23] to within the accuracy of one block.

The source has to send its data at the same time and in waves. After sending one wave of data, it has to wait long enough for that wave to propagate four blocks so as to ensure that the next wave of data will not interfere with the forwarding transmissions of the previous wave by the nodes in downstream blocks. Note that if the blocks are made small enough so that nodes on opposite ends of two adjacent blocks are within communication range of one another (e.g., nodes A and B), there will be nodes that are two blocks apart that will also be in each other's communication range (e.g., nodes C and D). Therefore, the source must wait an amount of time it takes for the wave to propagate four blocks before sending the next wave (e.g., the source must wait for the first wave to clear Block 3 before sending the second wave so that node C does not experience interference from Block 3's broadcast of the first wave). By having the source send a wave of data simultaneously and the intermediate nodes relay the wave as soon as they receive it, the network is made to operate as a time-slotted communication system.

Sending the data in waves prevents collisions among broadcasts from different blocks, however the fact that the transmitters are untuned results in collisions by transmissions within the same block. If every node in a block transmits on a random frequency, it is likely that there will be transmissions at frequencies close to each other, thus any receiver tuned to those frequencies will detect a collision and will not be able to decode the individual transmissions. These collisions effectively erase some of the packets, making it seem as if nodes in neighboring blocks communicate with each other through an erasure channel. The question we are interested in is, given N unit-capacity communication channels, how much data can simultaneously be sent to the destination and have this data successfully received and decoded by the destination that is H hops away. For now, let us assume that H is a constant, though later it will be shown that H may be allowed to grow with N, as long as $\log(H(N))/N$ goes to zero as N grows (i.e., as long as H grows at a rate that is less than exponential in N), without affecting the asymptotic throughput. We want to compare this to a fully-coordinated network employing tuned radios, in which case exactly N packets could be sent in each wave, provided that there are N active nodes in each block and every node selects a unique frequency on which to communicate (even in the case of perfectly tuned and coordinated radios, in order to avoid interference across waves, new waves must be sent four time slots apart). Note that, in order to have close to N nodes participating in communication in each block, either the nodes should be deployed with even density throughout the network or they should be adaptively duty cycled [20] so that only N nodes in each block are active.

IV. Throughput

We will find the relative throughput of the untuned network by showing that the connectivity of the network can be modeled as a random graph and then applying known results from network coding literature. Namely, we make use of the result that for communication in a graph of unit capacity links for which the connectivity is not known *a priori*, a throughput equal to the max-flow between the source and the destination is achievable with arbitrarily high probability by using random linear network coding[1] over a high enough field size [24]. However, in order to make use of this result, we must find the max-flow of the graph. This is done by Result 1. Since the connectivity of this random graph is not static (i.e., each wave of data will encounter a different set of links), the packets have to carry the encoding vectors in their headers to provide the destination with just the right information needed to decode the source packets as in the scheme introduced in [25].

We will show that the throughput with network coding is linear in N over $H(N)$ hops as long as $\log(H(N))/N$ goes to zero as N grows (i.e., as long as the growth of $H(N)$ is subexponential in N). In contrast, simple random routing (in which forwarding nodes are only allowed to randomly select one of the packets from each wave to forward, rather than combining the packets they receive in each wave to form the output packet) has constant throughput over a number of hops that is linear in N (i.e., the throughput per channel goes to zero as the number of available channels grows). This result is proven in Appendix B. In other words, random routing cannot provide throughput that grows with N if the number of hops, $H(N)$, grows at least linearly with N, while random network coding allows for throughput that scales linearly with N, over a number of hops, $H(N)$, that can grow with N, as long as the growth is less than exponential.

[1]In random linear network coding, each forwarding node sends on each outgoing link a random linear combination of the packets it receives on the input links. Each input packet is multiplied by a randomly chosen element from some Galois Field and these products are added together to form the outgoing packet [24].

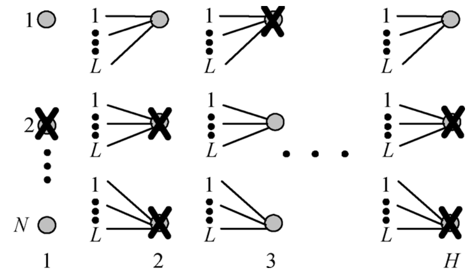

Fig. 3. Random graph representing connectivity in the network of nodes with untuned radios. Each vertex (node) in columns $\{2, \dots H\}$ has L inputs, each one coming from a randomly and uniformly selected node in the previous column. Each node, along with its incoming and outgoing links, is deleted with probability $1 - 1/e$.

A. Random Graph Representation

We now create the random graph, shown in Fig. 3, that models the connectivity of the network. The N vertices in each column correspond to the N nodes in each block during communication. The H columns correspond to communicating over H blocks (for ease of notation, we first consider the case when H is constant, rather than a function of N). Each of the vertices in columns $\{2, \dots, H\}$ has L incoming links corresponding to the L receivers on each node. Each link connects a vertex to a randomly, independently chosen vertex in the previous column. Since transmissions experience collisions with probability $1 - 1/e$, each of the vertices in the graph is deleted with probability $1 - 1/e$, corresponding to a collision on the channel on which the node transmits its packet.[2] When a vertex is deleted, all of its incoming and outgoing links are also deleted. This means that each of the links is deleted with probability $1 - 1/e$ because each receive-radio has probability $1 - 1/e$ of being tuned to a frequency range that does not contain a decodable transmission (i.e., either no transmission or more than one). The links that are not deleted are equally likely to connect to any of the vertices in the previous column that are not deleted because, given that a receive-radio has exactly one transmission in its receive frequency range, the source of that transmission is equally likely to be any of the transmitters that do not experience a collision.

We label the resulting random graph as $G_{L,1-1/e}$ and show that the max-flow of $G_{L,1-1/e}$ is close to N/e if L is large enough.

B. Max-Flow of Random Graph

Result 1: For any constant β such that $\beta < 1/e$ there exists a *constant* number of inputs/node L such that the max-flow of $G_{L,1-1/e}$ is greater than $\beta \cdot N$ with high probability as N goes to infinity.

The proof of Result 1 will proceed as follows.

- The first step is to relate the likelihood of having many disjoint end-to-end (E2E) paths to the likelihood of having at least one path E2E. This relationship is established in Lemma 1.

[2]In the random graph, the vertices are deleted independently of one another. This is not the case in the network since the collisions are not independent; however, this approximation becomes accurate as N tends to infinity. The independence assumption allows for analytical tractability in what follows.

- The second step is to derive an upper bound on the probability that no paths exist E2E. This is done in Lemma 2.
- Finally, using the relationship established in Lemma 1 and the upper bound derived in Lemma 2, we can find an upper bound on the probability not having $\beta \cdot N$ paths E2E, and show that it decays exponentially with N.

We define the following notation. Let A be the set of graphs that have at least one path E2E and let A_r be the set of graphs having the following property: starting with any graph in A_r, the deletion of *any* r vertices will still result in a graph belonging to the set A. This is equivalent to saying that any graph in A_r has at least $r+1$ vertex-disjoint paths E2E.

Lemma 1: Let r be a positive integer. Then

$$P(G_{L,p_2} \notin A_r) \leq \left(\frac{q_2}{q_2 - q_1}\right)^r P(G_{L,p_1} \notin A) \quad (1)$$

for each $0 \leq p_2 \leq p_1 \leq 1$. Here, $q_1 = 1 - p_1$ and $q_2 = 1 - p_2$, and $G_{L,p}$ is a random graph in which each vertex originally has L incoming links and the vertices are deleted with probability p, as described in Section IV-A.

Proof of Lemma 1: This proof is based on the proof in [26] of a similar result from Percolation Theory. Let $X_{i,j}\ \forall i \in \{1,\ldots,N\}$ and $j \in \{1,\ldots,H\}$ be i.i.d. random variables uniformly distributed in the interval [0, 1], and to each vertex in row i and column j of the grid assign the value $X_{i,j}$. To create graphs G_{L,p_1} and G_{L,p_2} that have vertices deleted with probability p_1 and p_2 respectively, do the following: first assign the values $X_{i,j}$ to each vertex in the grid. Then assign L links from each node in columns 2 through H to a randomly selected node in the previous column. Finally, to create graph G_{L,p_1}, for each vertex i, j, delete it iff $X_{i,j} \leq p_1$. To create graph G_{L,p_2}, for each vertex i, j, delete it iff $X_{i,j} \leq p_2$.

We are interested in relating the likelihood that $G_{L,p_2} \in A_r$ to the likelihood that $G_{L,p_1} \in A$. Note that if $G_{L,p_2} \notin A_r$, then there must be a set of vertices, B, such that

a) aone of the vertices in the set B are deleted in G_{L,p_2};
b) $|B| \leq r$;
c) the graph $\overline{G}_{L,p_2}$ obtained by deleting from G_{L,p_2} the vertices in B satisfies $\overline{G}_{L,p_2} \notin A$.

There may exist many such sets B, in which case it is sufficient to pick any such set. Suppose that $G_{L,p_2} \notin A_r$, and that every vertex i, j in the set B satisfies $p_2 < X_{i,j} \leq p_1$. It then follows from c) that $G_{L,p_1} \notin A$. Conditional on B, there is a $[(p_1 - p_2)/(1-p_2)]^{|B|} = [(q_2 - q_1)/q_2]^{|B|}$ probability that $p_2 < X_{i,j} \leq p_1$ for all vertices in B; therefore

$$P(G_{L,p_1} \notin A | G_{L,p_2} \notin A_r) \geq \left(\frac{q_2 - q_1}{q_2}\right)^r. \quad (2)$$

Applying Bayes's theorem and the fact that

$$P(G_{L,p_1} \notin A \cap G_{L,p_2} \notin A_r) \leq P(G_{L,p_1} \notin A)$$

gives the result of Lemma 1. □

This result of Lemma 1 is particularly useful if we can show that the probability that that $G_{L,p_1} \notin A$ decays exponentially (with N) to zero for some p_1. In other words, we wish to show that $P(G_{L,p_1} \notin A) \leq e^{-N\alpha(p_1,L)}$. We are also interested in the rate of this exponential decay, $\alpha(p_1, L)$, which depends on both p_1 and L. By showing that $P(G_{L,p_1} \notin A) \leq e^{-N\alpha(p_1,L)}$, we can rewrite the result of Lemma 1 as

$$P(G_{L,p_2} \notin A_r) \leq \left(\frac{q_2}{q_2 - q_1}\right)^r e^{-N\alpha(p_1,L)} \quad (3)$$

and apply $r = \beta \cdot N$ to show that the probability of not having $\beta \cdot N$ (actually $\beta \cdot N + 1$) paths decays to zero exponentially as long as

$$\beta < \alpha(p_1, L) / \log\left(\frac{q_2}{q_2 - q_1}\right). \quad (4)$$

The problem now becomes finding an appropriate bound on $P(G_{L,p_1} \notin A)$.

Lemma 2: If $L(1-p_1) > 1$, then

$$P(G_{L,p_1} \notin A) \leq H \cdot P(Y < (1 - Z^*)N)$$

where Y is a random variable drawn from the

$$\text{Binomial}\left(N, 1 - p_1 + (1-p_1) Z^{*L}\right)$$

distribution and $Z^* = [L(1-p_1)]^{1/(1-L)}$.

Lemma 2 allows us to relate the probability that no E2E path exists in G_{L,p_1} to the probability that a binomially distributed random variable deviates from its expected value by an amount proportional to N. Since this probability decays to zero exponentially in N, this result, along with Lemma 1 will allow us to prove that the number of vertex-disjoint paths in G_{L,p_2} will be linear in N with high probability. The constant, β, of this linear relationship will depend on the value of L. Note that β also depends on the value p_1; however, the value p_1 is not fundamental to the graph G_{L,p_2}, and we are allowed to assign any value to p_1, as long as it is larger than p_2, so as to maximize the bound on β guaranteed by Lemma 1 and Lemma 2.

Also note that the condition that $L(1-p_1) > 1$ is imposed to ensure that $1 - Z^* > 0$. It can be shown that if $L(1-p_1) > 1$, then $G_{L,p_1} \in A$ with high probability, otherwise $G_{L,p_1} \notin A$ with high probability. However, in our case it is not enough to show that $G_{L,p_1} \in A$ with high probability for appropriate values of L and p_1. We must also bound the rate of this convergence, using Lemma 2, in order to apply Lemma 1 to our original problem (showing that the max flow of G_{L,p_2} grows linearly with N).

The proof of Lemma 2 will proceed as follows.

- Consider the number of vertices in each column of G_{L,p_1} that were not deleted and have a path back to column 1. We call these "good" vertices.
- We set a threshold and consider the probability that the number of good vertices in a column ever falls below this threshold. This probability is an upper bound on the probability of having no paths end to end.
- We show that the probability that the number of good vertices in column j falls below the threshold, given that the number of good vertices in column $j-1$ is on or above the threshold, is upper bounded by the probability that the number of good vertices in column j falls below the threshold, given that the number of good vertices in column $j-1$ is exactly on the threshold.

- This probability is then shown to be equal to the probability that a binomially distributed random variable deviates from its mean by some amount proportional to N.

Proof of Lemma 2: Consider the number of vertices in each column of G_{L,p_1} that were not deleted and have a path back to column 1. Let us call these vertices "good" and all the others "bad." Conditioned on the number of bad (good) vertices in column j, vertices in column $j+1$ are themselves good or bad independently of each other and with equal probability. Let the number of bad vertices in column j be $Z \cdot N$ for some Z that satisfies $0 \leq Z \leq 1$. Then vertices in column $j+1$ are themselves bad with probability $p_1 + (1-p_1)\,Z^L$.

Consider the probability that the number of bad nodes in column $j+1$ is greater than $Z \cdot N$, given that the number of bad nodes in column j is $Z \cdot N$. If $p_1 + (1-p_1)\,Z^L < Z$, this probability should be exponentially small in N. This probability is minimized when the difference $Z - \left(p_1 + (1-p_1)\,Z^L\right)$ is greatest. Setting the first derivative of this function, which is convex in the interval $[0, 1)$, to zero shows that the expression is maximized when $Z = [L\,(1-p_1)]^{1/(1-L)}$ – we label this value as Z^*. We prove Lemma 2 by arguing that $P\,(G_{L,p_1} \notin A)$ is upper bounded by the probability that the number of bad vertices in any column is greater than Z^*, given that the number of bad vertices in column 1 is less than $Z^* N$.

Let R_j represent the ratio of bad nodes to the total number of nodes in column j (i.e., the total number of bad nodes in column j is $R_j \cdot N$). $G_{L,p_1} \notin A$ is equivalent to $R_H = 1$. Note that if $R_j = 1$ for some $j \in [1, H)$, then $R_k = 1\, \forall k \in [j+1, H]$ because if column j has no connectivity back to column 1, then none of the columns after j will have connectivity to column 1 either. We prove Lemma 2 by arguing that $P\,(G_{L,p_1} \notin A)$ is equal to the probability that there exists a $j \in [1, H]$ for which $R_j = 1$, and this is upper bounded by the probability that there exists a $j \in [1, H]$ for which $R_j > Z^*$. Mathematically

$$\begin{aligned}
&P\,(G_{L,p_1} \notin A)\\
&\quad < P_{G_{L,p_1}}\,(\exists j \in [1, H] \text{ s.t. } R_j > Z^*)\\
&\quad \overset{(a)}{\leq} P_{G_{L,p_{1\,1}}}\,(R_1 > Z^*)\\
&\qquad + \sum_{j=2}^{H} P_{G_{L,p_1}(R_j > Z^* | R_{j-1} \leq Z^*)}\\
&\quad < P_{G_{L,p_1}}\,(R_1 > Z^*)\\
&\qquad + \sum_{j=2}^{H} P_{G_{L,p_1}(R_j > Z^* | R_{j-1} = Z^*)}\\
&\quad \overset{(b)}{=} P_{G_{L,p_1}}\,(R_1 > Z^*) + (H-1) \cdot P(Y < (1-Z^*)N)\\
&\quad \overset{(c)}{<} H \cdot P\,(Y < (1-Z^*)N)
\end{aligned} \tag{5}$$

where Y is a random variable drawn from the

$$\text{Binomial}\left(N, 1 - \left[p_1 + (1-p_1)\,{Z^*}^L\right]\right)$$

distribution. (a) is due to the union bound. (b) is because, given $R_{j-1} = Z^*$, each vertex in column j is bad with probability $p_1 + (1-p_1)\,{Z^*}^L$. (c) is because each vertex in column 1 is bad with probability p_1, which is less than the probability that a vertex in column j is bad, conditioned on $R_{j-1} = Z^*$. □

Proof of Result 1: Now, we must find the rate at which

$$P\,(Y < (1-Z^*)N) = P\,(Y/N < 1 - Z^*)$$

decays to zero and apply this to Lemma 1 to show that the number of vertex-disjoint paths in G_{L,p_2} grows linearly with N. Applying the Chernoff bound [27] to $P\,(Y/N < 1 - Z^*)$ gives us an appropriate upper bound on this rate. We use the following notation: $q = 1 - [p_1 + (1-p_1)\,Z^{*L}]$ and $\varepsilon = q - 1 + Z^*$ to write [27]:

$$\begin{aligned}
&P\,(Y/N < 1 - Z^*)\\
&\quad = P\,(q - Y/N > \varepsilon) \leq \left(\frac{q-\varepsilon}{q}\right)^{-N(q-\varepsilon)} \left(\frac{1-q+\varepsilon}{1-q}\right)^{-N(1-q+\varepsilon)}.
\end{aligned} \tag{6}$$

The right-hand side of the equation can also be expressed as

$$\exp\left(\log\left[\left(\frac{q-\varepsilon}{q}\right)^{-N(q-\varepsilon)} \left(\frac{1-q+\varepsilon}{1-q}\right)^{-N(1-q+\varepsilon)}\right]\right) \tag{7}$$

giving us

$$\alpha\,(p_1, L) = \log\left[\left(\frac{q-\varepsilon}{q}\right)^{(q-\varepsilon)} \left(\frac{1-q+\varepsilon}{1-q}\right)^{(1-q+\varepsilon)}\right] \tag{8}$$

as the $\alpha\,(p_1, L)$ we need to plug into (4). Evaluating (4) with a large enough *but constant* value for L and appropriately chosen value for p_1 provides the guarantee that the max-flow of $G_{L,1-1/e}$ is at least $\beta \cdot N$ for any $\beta < 1/e$, proving Result 1. □

The throughput guaranteed by Result 1 is not dependent on H because we held H constant while letting N go to infinity. However, (5) implies that H may grow with N without affecting the throughput result as long as $\log(H(N))/N$ goes to zero as N grows (i.e., as long as the rate of growth of H is less than exponential in N), because all we need is for $P\,(G_{L,p_1} \notin A)$ to decay exponentially with N at such a rate that the right-hand side of (3) decays exponentially with N. Since $P\,(Y/N < 1 - Z^*)$ decays exponentially with N, (5) tells us that so does $P\,(G_{L,p_1} \notin A)$ as long as $\log(H(N))/N$ goes to zero as N grows. Even if H grows exponentially with N, the throughput would still be linear in N, as long as the rate of this exponential growth were less than $1/e$ (i.e., as long as $\log(H(N))/N < 1/e$). The exponential growth of $H(N)$ would just reduce the constant of the linear throughput of the network by reducing the value of $\alpha\,(p_1, L)$ in (3) and (4). We can restate this result as a corollary to Result 1:

Corollary 1: For any constant β such that $\beta < 1/e$, and number of hops $H(N)$ such that $\beta > \log(H(N))/N$, there exists a *constant* number of inputs/node L such that the max-flow of $G_{L,1-1/e}$ is greater than $N \cdot (\beta - \log(H(N))/N)$ with high probability as N goes to infinity.

Note that the Result 1 is proven using bounding techniques, so it is not tight. For example, for $L = 10$, the result proves that the throughput is guaranteed to be at least 2%, whereas simulations show that $L = 10$ is good enough to give throughput of 30%. Fig. 4 shows the simulation results. A network with $N = 1000$ nodes per block and $H = 1000$ blocks was simulated. The figure shows the average throughput at various values of inputs/node, L. One thousand simulations were run for each value of L.

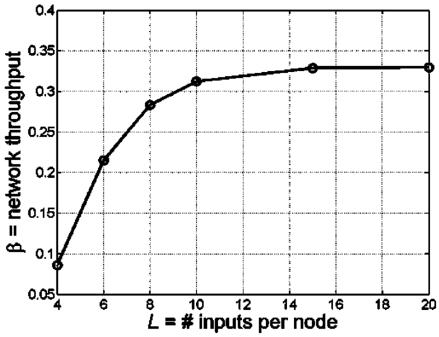

Fig. 4. Simulation results showing the ratio of the throughput achievable with untuned radios and network coding to the throughput of a perfectly tuned and synchronized network, as a function of the number of inputs per node for 1000 channels and 1000 nodes per block.

This max-flow result, together with the random network coding result of [24] tells us that each wave of data can deliver nearly N/e packets from the sources to the destination, compared to N packets that could be transported in each wave if the network were composed of nodes with tuned radios and perfect coordination.

In practical deployments, having about 10 radios per node would be realistic. In this case, the theoretical results tell us that a throughput of only $0.02 \cdot N$ can be guaranteed. However, simulations show that a throughput of $0.3 \cdot N$ can be expected. This is a good trade-off for many applications in which the demand on bandwidth is not as strict as the demand for low-cost nodes that can operate at power levels comparable to the power levels that can be supplied by the energy scavenging mechanisms.

V. Conclusion

This paper shows how network coding can be used to achieve high throughput in an *ad-hoc* wireless network of nodes with untuned radios. This coding technique makes it possible to build ultralow-cost and low-power devices, and deploy a high-performance network of unreliable nodes by utilizing randomized algorithms and high density of such nodes.

Appendix A

Consider placing k balls into n bins such that each ball randomly selects a bin in which to be placed independently of all other balls, and all of the random selections are made using the same probability mass function (pmf). Given the number of bins, n, we wish to find the number of balls, k, and the pmf such that the expected number of bins containing exactly one ball is maximized.

Result 2: The expected number of bins containing exactly one ball is maximized when the number of balls is equal to the number of bins and each ball is equally likely to select any of the bins. In this case, the expected number of bins containing exactly one ball, in the limit as n grows to infinity, is n/e.

Proof: We first write the expression for the expected number of bins containing exactly one ball using b_i to denote the event that bin i contains exactly one ball, p_i to denote the probability that any given ball chooses bin i, and $1\{\cdot\}$ to denote the indicator function, which takes on the value 1 if the expression in the braces is true and the value 0 otherwise.

$$\begin{aligned} E\left[\sum_{i=1}^{n} 1\{b_i\}\right] &= \sum_{i=1}^{n} E\left[1\{b_i\}\right] \\ &= \sum_{i=1}^{n} P\left[b_i\right] \\ &= \sum_{i=1}^{n} \binom{k}{1} p_i \left(1-p_i\right)^{k-1} \\ &= k \sum_{i=1}^{n} p_i \left(1-p_i\right)^{k-1}. \end{aligned} \tag{9}$$

Therefore, we seek to find the

$$\underset{p_i \text{s.t.} \sum_{i=1}^{n} p_i = 1}{\arg\max} \sum_{i=1}^{n} p_i \left(1-p_i\right)^{k-1}. \tag{10}$$

Let us first find the value of p_i that maximizes each of the summands in (10). Solving for:

$$0 = \frac{d}{dp} p\left(1-p\right)^{k-1} = \left(1-p\right)^{k-1} - (k-1) \cdot p \cdot \left(1-p\right)^{k-2} \tag{11}$$

gives $1/k$ as a local maximum of the summands in (10). This value also gives the global maximum in the range $[0,1]$ because the function has a strictly positive first derivative in the range $[0, 1/k)$ and a strictly negative first derivative in the range $(1/k, 1)$.

Since the value of p_i in the range $[0,1]$ that individually maximizes each of the summands of (10) is $1/k$, the overall sum is maximized when $p_i = 1/k \, \forall i \in \{1, \ldots, n\}$ (i.e., each of the balls is equally likely to be in any of the bins). However, due to the constraint that the p_i's form a valid pmf (i.e., $\sum_{i=1}^{n} p_i = 1$), this is only possible when $k = n$. Therefore, the expected number of bins containing exactly one ball is maximized when the number of balls is equal to the number of bins and each ball is equally likely to be placed in any of the bins.

The expected number of bins containing exactly one ball can be computed as

$$\begin{aligned} k \sum_{i=1}^{n} p_i \left(1-p_i\right)^{k-1} &= n \sum_{i=1}^{n} p_i \left(1-p_i\right)^{n-1} \\ &= n \sum_{i=1}^{n} (1/n)\left(1-1/n\right)^{n-1} \\ &= n \cdot \left(1-1/n\right)^{n-1}. \end{aligned} \tag{12}$$

Since $\lim_{n\to\infty} \left(1-1/n\right)^{n} = 1/e$, the expression in (12) asymptotically goes to n/e as n grows large. □

Appendix B

To analyze the throughput of a routing scheme in which the output of each intermediate node is simply a copy of one of its inputs (the node randomly selects which of the inputs to forward), rather than a function of all of the inputs, consider the random graph that corresponds to this routing scheme. The

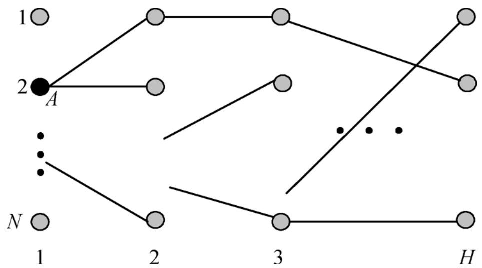

Fig. 5. Random graph representing connectivity when only routing is allowed. Each vertex in columns $\{2,\ldots,H\}$ connects to one randomly chosen vertex in the previous column.

TABLE I
EXPECTED THROUGHPUT OF ROUTING OVER H STEPS

N	$H=N$	$H=10N$
100	2.304	1.0001
400	2.351	1.0001
700	2.359	1.0001
1000	2.362	1.0001
2000	2.362	1.0001

The throughput over N steps remains constant, even as N grows. Over $10N$ steps, only one packet can be sent end-to end with routing, even if no collisions occur.

connectivity of the network is still the same and can be represented by Fig. 3; however, because each intermediate node only forward a packet from a randomly selected input link, the graph must be modified to reflect this. Because only one of the incoming packets is forwarded, the information coming in on the other links is not propagated; therefore, those links can be deleted from the graph without altering the throughput. In other words, starting with the graph in Fig. 3, for each vertex that is not deleted and has at least one incoming link (i.e., at least one of the original L incoming links was connected to a surviving vertex in the previous column), randomly chose one of those incoming links to keep and delete all the others. Which link is kept at each vertex is chosen uniformly and independently from the other vertices. The problem then becomes finding the end to end max-flow of the resulting graph.

To make the problem more analytically tractable, we consider the bounding case in which none of the vertices in the graph are deleted. In this case, the random graph simply consists of vertices in an $N \times H$ grid with each vertex in columns $\{2,\ldots,H\}$ connecting to only one randomly selected (with uniform probability) vertex in the previous column. An instance of this random graph is shown in Fig. 5. Since deleting vertices (along with the links associated with those vertices) can only decrease the max-flow of a graph, any upper bound on the max-flow from column 1 to column H is also an upper bound on the max-flow of the same graph in which certain vertices are deleted (remember that the performance of a routing scheme in which only forwarding is allowed corresponds to the max-flow of the graph in Fig. 5 with each vertex deleted with probability $1-1/e$).

To find the max-flow of this graph, consider a particular vertex in Column 1. Let us label this vertex as vertex A, as shown in Fig. 5. Consider the number of vertices in columns $\{2,\ldots,H\}$ that have a connection back to vertex A (note that, since each vertex has only one incoming link, each one has a connection to only one vertex in the first column). Since each vertex randomly and independently connects to a vertex in the previous column, the number of vertices in each column that connect back to vertex A only depends on the number of vertices in the previous column that connect to A; therefore, the number of vertices in each column with a connection back to A forms a Markov chain. The transition probability matrix P of this Markov chain has the form

$$P_{i,j} = \binom{N}{j}\left(\frac{i}{N}\right)^{j}\left(\frac{N-i}{N}\right)^{N-j} \tag{13}$$

for $i,j \in \{0,\ldots,N\}$. Here $P_{i,j}$ gives the probability that j vertices of a column have a connection back to A given that i vertices in the previous column have a connection back to A.

This Markov chain has two absorbing states 0 and N, while all the others are transient. Therefore, the chain is guaranteed eventually to be absorbed in one of these two states. Since the number of descendents of each vertex in column 1 forms a Markov chain with the same transition probability matrix and the number of descendents of each of these vertices has to add up to N, eventually all but one of the chains will be absorbed in the zero state, while one of the chains will be absorbed in state N. In other words, eventually there will be a column in which all of the vertices connect back to the same vertex of column 1, and all subsequent columns will likewise only connect back to that vertex.

Indeed, row 1 (corresponding to starting in state 1, with only one vertex connecting back to A) of P^{∞} has the form $[1-1/N, 0, \ldots, 0, 1/N]$. $P^{\infty}_{i,j}$ gives the probability of starting in state i and ending in state j after infinitely many steps. Row 1 of P^{∞} shows that any vertex in column 1 will eventually either have no descendents (this happens with probability $1-1/N$) or all of the vertices in later columns will be its descendents (this happens with probability $1/N$). This means that, as H goes to infinity, the max-flow of the graph will be 1.

But what happens over a finite number of steps? The expected max-flow over H steps is given by:

$$N \cdot P(\text{given vertex has descendents after } H \text{ steps} = N \cdot \left(1 - P^{H}_{1,0}\right). \tag{14}$$

The problem is that the closed-form expression for P^{H} is unmanageably elaborate; therefore, we rely on a computer to evaluate it for some values of N and H. Table I shows the expected max-flow between column 1 and column H for various values of N in the graph corresponding to a scheme that only allows routing at intermediate nodes and in the case when no collisions occur (i.e., no vertices are deleted from the graph). As can be seen from the table, the performance of a routing-only scheme allows for a constant throughput over N hops, while network coding allows for throughput that is linear in N over H hops as long as $\log(H(N))/N$ goes to zero as N grows (i.e., as long as the rate of growth of H is less than exponential in N).

ACKNOWLEDGMENT

The authors would like to thank Brian Otis, Massimo Franceschetti, Alistair Sinclair, Vinod Prabhakaran, and Alvise

Bonivento for sharing their expertise and insights during the course of this work.

REFERENCES

[1] [Online]. Available: http://www.tinyos.net/

[2] [Online]. Available: http://webs.cs.berkeley.edu/nest-index.html

[3] [Online]. Available: http://bwrc.eecs.berkeley.edu/Research/Pico_Radio/Default.htm

[4] [Online]. Available: http://www.cens.ucla.edu/

[5] [Online]. Available: http://www.intel-research.net/berkeley/index.asp

[6] [Online]. Available: http://www.zigbee.org

[7] [Online]. Available: http://www.xbow.com

[8] [Online]. Available: http://www.ember.com

[9] [Online]. Available: http://www.dust-inc.com/flash-index.shtml

[10] [Online]. Available: http://www.cs.berkeley.edu/~binetude/ggb/

[11] [Online]. Available: http://www.cbe.berkeley.edu/research/briefs-Wireless.htm

[12] "Intel Res.,", IRB-TR-02-006, Jun. 19, 2002.

[13] S. Coleri, S. Cheung, and P. Varaiya, "Sensor networks for monitoring traffic," in *Proc. 42nd Ann. Allerton Conf. Commun., Contr. Comput.*, Sep. 2004.

[14]

[15] S. Roundy, B. Otis, Y. H. Chee, J. Rabaey, and P. Wright, "A 1.9 GHz RF transmit beacon using environmentally scavenged energy," in *Proc. Dig. IEEE Int. Symp. Low Power Elec. Devices*, Seoul, Korea, 2003.

[16] B. Otis, Y. H. Chee, R. Lu, N. Pletcher, and J. Rabaey, "An ultra-low power MEMS-based two-channel transceiver for wireless sensor networks," in *Proc. IEEE Symp. VLSI Circuits*, Honolulu, HI, Jun. 2004.

[17] A. Willig *et al.*, "Measurements of a wireless link in an industrial environment using an IEEE 802.11-Compliant physical layer," *IEEE Trans. Ind. Electron.*, vol. 43, Dec. 2002.

[18] N. Patwari, Y. Wang, and R. O'Dea, "The importance of the multipoint-to-multipoint indoor radio channel in ad hoc networks," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Orlando, FL, Mar. 2002.

[19] M. Zorzi and R. R. Rao, "Geographic random forwarding (GeRaF) for ad hoc and sensor networks: Energy and latency performance," *IEEE Trans. Mobile Comput.*, vol. 2, Oct.–Dec. 2003.

[20] J. Van Greuen, D. Petrović, A. Bonivento, J. Rabaey, K. Ramchandran, and A. Sangiovanni-Vincentelli, "Adaptive sleep discipline for energy conservation and robustness in dense sensor networks," in *Proc. ICC 2004*.

[21] Y. Cheng, "The influence and modeling of process variation and device mismatch for analog/RF circuit design," in *Proc. IEEE Conf. Devices, Circuits Syst.*, 2002.

[22] C. P. Yue and S. S. Wong, "On-chip spiral inductors with patterned ground shields for Si-based RF IC's," *IEEE J. Solid-State Circuits*, vol. 33, May 1998.

[23] L. Doherty, L. El Ghaoui, and K. Pister, "Convex position estimation in wireless sensor networks," in *Proc. IEEE INFOCOM 2001*, Anchorage, AK, Apr. 2001.

[24] T. Ho, M. Medard, J. Shi, M. Effros, and D. Krager, "On randomized network coding," in *Proc. 41st Ann. Allerton Conf. Commun., Contr., Comput.*, Oct. 2003.

[25] P. Chou, Y. Wu, and K. Jain, "Practical network coding," in *Proc. Allerton Conf. Commun., Contr., Comput.*, Monticello, IL, Oct. 2003.

[26] M. Aizenman, J. T. Chayes, L. Chayes, J. Fröhlich, and L. Russo, "On a sharp transition from area law to perimeter law in a system of random surfaces," *Commun. Math. Phys.*, vol. 92, pp. 19–69, 1983.

[27] H. Chernoff, "A measure of asymptotic efficiency for test of a hypothesis based on the sum of observations," *Ann. Math. Stat.*, vol. 23, pp. 493–507, 1952.