# Bounds on the Network Coding Capacity for Wireless Random Networks

Salah A. Aly, Vishal Kapoor, Jie Meng, Andreas Klappenecker
*Department of Computer Science,*
*Texas A&M University*
*College Station, TX - 77840*
{salah, vishal, jmeng, klappi}@cs.tamu.edu

***Abstract*— Recently, it has been shown that the max flow capacity can be achieved in a multicast network using network coding. In this paper, we propose and analyze a more realistic model for wireless random networks. We prove that the capacity of network coding for this model is concentrated around the expected value of its minimum cut. Furthermore, we establish upper and lower bounds for wireless nodes using Chernoff bounds. Our experiments show that our theoretical predictions are well matched by simulation results.**

## I. Introduction

Traditionally, the information flow in networks is modeled as a multi-commodity flow problem by treating the underlying network as a flow network. Suppose that one source node in a graph has to transfer some information to one destination node (i.e., a unicast situation). By Menger's theorem [3], the maximum information that can flow is upper bounded by the value of the minimum cut between the source and the destination; this well-known result from classical graph theory is also known as the *Max-flow Min-Cut theorem*. One can use max-flow min-cut algorithms to compute the maximum throughput for instance for unicast, multicast, and multi-source multicast communications.

Recently, Ahlswede, Cai, Li, and Yeung proposed in the seminal paper [1] a new paradigm, called *network coding*. Their key observation was that traditional store-and-forward networks cannot always achieve the max-flow value, whereas one can achieve this value using network coding. The idea is based on the simple fact that information can be replicated, mixed together and then transmitted over links to save bandwidth. If this is properly done, then the information can be reliably decoded at the receiver nodes, see e.g. [1], [16]. The basic idea of network coding is that the intermediate network nodes can now process, encode, and transmit information.

Since its inception by Ahlswede et al., there has been an upsurge of interest in network coding, see for example [4]–[8], [10], [12] and the references therein. Arguably, most network coding publications model the underlying network as a directed acyclic graph and are typically concerned with solving single source multicast or multi-source multicast using deterministic or randomized encoding and decoding schemes.

In this paper, we discuss a new model for wireless random networks. In this model, nodes are placed at random locations. Two nodes $u$ and $v$ are connected with probability 1 if the distance between them is less than $r$; the nodes are connected with probability $p < 1$ if the distance between them is less than or equal to $R$ but greater than $r$; otherwise $u$ and $v$ are not connected. Thus, the model is a refinement of geometric random graphs that incorporates the potential loss of connectivity towards the end of the transmission range, where interference is more dominant. The main contributions of this paper are:

- We introduce the quasi random geometric graph model, a model of wireless network topologies that simulates the connectivity in mobile ad-hoc networks more realistically than the random graph model, but is still easy to analyze.
- We derive high-probability bounds for the network coding capacity of quasi random geometric graphs.
- We provide simulations results that support our bounds on the network coding capacity.

The rest of this paper is organized as follows. In Section II, we give an overview of network coding and the previous work in capacity of network coding. In Section III, we present our new model. We provide our main results in Sections IV and V.

## II. Background and Model Description

In this section, we give a short summary of network coding, focusing on the calculation of the capacity of a min cut in a weighted random graph. For a more in depth discussion of basic concepts and methods of network coding, we refer the reader to the survey paper [4].

### A. Network Coding Fundamentals

To illustrate the power of network coding, we provide a simple example, which is often referred to as the *Wheatstone bridge*, due to its electrical circuits origin. It demonstrates that multicast routing can achieve the maximum possible throughput in a communication network using a coding scheme consisting of linear operations in finite field, whereas traditional store-and-forward routing cannot achieve the same throughput.

Consider the example shown in Fig.1(a), where the nodes $X$ and $Y$ respectively want to send two bits $b_1$ and $b_2$ to each other. One way of doing this is to let the bit $b_1$ travel on the path $X \rightarrow A \rightarrow B \rightarrow Y$ at one point of time and to let $b_2$ travel on the path $Y \rightarrow A \rightarrow B \rightarrow X$ on the other. However, if the network wants to transmit the bits simultaneously, then there is no way to do so, as there are no disjoint paths between $X$ and $Y$.

However, using network coding as shown in Fig. 1(b), one can save bandwidth. In this case, both $X$ and $Y$ transmit the bits $b_1$ and $b_2$ (as shown in the figure) and then $A$ XORs (*encodes*) them together and the resulting bit $b_1 \oplus b_2$ travels over the paths $A \to B \to Y$ and $A \to B \to X$. Since node $X$ already has of $b_1$, it can recover (*decode*) $b_2$ by the operation $b_1 \oplus (b_1 \oplus b_2)$. Similarly $Y$ can also decode $b_1$.

This example illustrates that the capacity of the minimum cut (equal to 1 in this example) can be easily achieved by network coding, whereas two rounds are needed to achieve the multicast in the uncoded (traditional) routing case, assuming unit capacity edges. Because of such benefits, network coding can be used in wireless ad-hoc networks or sensor networks to help conserve energy and to increase the overall throughput.

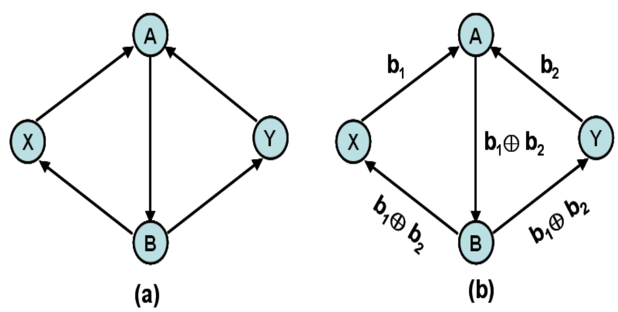


Fig. 1. An example of network coding on a Wheatstone Bridge

### B. Network Coding in Ad-hoc Wireless Networks

In [15], Ramamoorthy et al. modeled the capacities of the connected edges in a wireless network as a Weighted Random Geometric Graph ($\mathcal{G}^{WRGG}$) and considered the *single source multicast problem.*

*Definition 1 (Single Source Multicast Problem):* Let $G = (V, E)$ be a graph with vertex set $V$ and edge set $E$ representing a network. Let $S \subseteq V$ be a set of sources (origins) and $T \subseteq V$ be a set of terminals (destinations). The multicast problem is to distribute the messages from the senders $s \in S$ to all terminal nodes $t \in T$, allowing routing along the edges of $G$. In network coding, the vertices are allowed to encode the incoming bits (or packets) and send encoded versions along the outgoing edges. A *single source multicast problem* is the special case where one has a single sender, that is, $|S| = 1$.

Ramamoorthy et al. extended the results proved by Karger et al. in [9] and used them to derive bounds for coding capacity for a single source multicast problem in a network comprised of a single source $s$, an intermediate network consisting of $n$ relay nodes, and $l$ terminal nodes, having independent and identically distributed link capacities $\sim X$ between any two nodes. They showed that the network coding capacity is concentrated around the value $n\mathbf{E}[X]$ in such a network.

In this paper, we extend their work to a more general and more realistic model that we call the *Quasi Random Geometric Graph* model ($\mathcal{G}^{QRGG}$). We derive high-probability bounds for the network coding capacity of such graphs.

## III. Modeling Random Wireless Networks

In this section, we present our new model and study the capacity of a minimum cut in a random wireless network.

Let $r$ be a real number in the range $0 \leq r \leq 1$. Recall that a Random Geometric Graph is a graph $\mathcal{G}^{RGG} = (V, E)$ with $n$ nodes selected independently and uniformly at random from the unit square $[0, 1]^2$ in which any two nodes $u$ and $v$ in $V$ are connected by an edge $(u, v)$ in $E$ if and only if the Euclidean distance $d(u, v) \leq r$. Such a graph is rough approximation of wireless networks.

Random geometric graphs have been popular in wireless mobile ad-hoc networks literature, since it is a theoretical model of the network topology that is easy to analyze. However, it does not realistically model the area of transmission, which is, in general, not a disk of radius $r$. Recently, a more realistic model for connectivity was proposed by Kuhn, Wattenhofer, and Zollinger [11]. In their model, two nodes $u$ and $v$ may or may not be connected when their Euclidean distance $d(u, v)$ is within the range $r < d(u, v) \leq r'$, see Fig. 2. We use random instances of such quasi-disk graphs to model the dynamically changing network topology in wireless random ad-hoc networks.

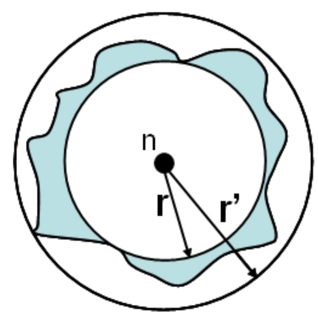


Fig. 2. The transmission range and Quasi Disk Graph Representation for a node

*Definition 2 (Quasi Random Geometric Graph ($\mathcal{G}^{QRGG}$)):* Let $r$ and $r'$ be two real numbers in the range $0 \leq r < r' \leq 1$. Let $V$ be a set of $n$ nodes that are selected independently and uniformly at random from the unit square $[0, 1]^2$. If $u$ and $v$ are two nodes in $V$, then

1) $(u, v) \in E$ if $d(u, v) \leq r$;
2) $(u, v) \notin E$ if $d(u, v) > r'$;
3) $(u, v) \in E$ with probability $p$ if $r < d(u, v) \leq r'$.

We call $\mathcal{G}^{QRGG} = (V, E)$ a quasi random geometric graph.

The difference between quasi random geometric graphs and random geometric graphs is that nodes at distance $d$ within the range $r < d \leq r'$ may or may not be connected; this models the connectivity in a more realistic way.

*Remark 3:* Instead of having a fixed probability $p$ for the connectivity of nodes within distance $d$ in the range $r < d \leq r'$, one can use a function $p(d)$ that associates a probability that depends on the distance to model the attenuation of the signal. Such a change is of course straightforward. We give one example in Section V.

In this paper, we consider the problem of single-source multicasts in such quasi random geometric graphs. Our main concern is to provide a lower bound for the capacity of network coding in this situation. Before defining the capacity, we need to further detail our model of connectivity.

*Definition 4 (Connectivity Graph):* Let $s$ be a source node, $T$ a set of terminal nodes, and $R$ a set of relay nodes. We define a connectivity graph $G = (V, E)$ as a graph with vertex set $V = \{s\} \cup R \cup T$ such that $G \in \mathcal{G}^{QRGG}$; in particular, the vertices are located in a unit square. We assume further that the source node only sends messages and terminal nodes only receive messages; in particular, the source and terminal nodes do not relay any messages. Furthermore, we assume that the source and the terminal nodes do not communicate directly; thus, any message is routed through at least one relay node.

We assume that the edges in the connectivity graph represent links with unit capacity. Put differently, we assume that the capacity $C_{ij}$ for $i, j$ in $V$ is given by

$$C_{ij} = \begin{cases} 1 & \text{if } (i,j) \in E, \\ 0 & \text{otherwise.} \end{cases}$$

We note that $C_{ij} = C_{ji}$, since the graph is undirected.

*Definition 5 (A Cut and its Capacity):* Let $G = (V, E)$ be a connectivity graph with source node $s$, a set $T$ of terminal nodes, and a set $R$ of relay nodes such that $V = \{s\} \cup R \cup T$. Let $t$ be a terminal node in $T$. An $s$-$t$-cut of size $k$ in the connectivity graph $G$ is a partition of the set of relay nodes $R$ into two sets $V_k$ and $\overline{V}_k$ such that

(i) $|V_k| = k$ and $|\overline{V}_k| = n - k$;
(ii) $R = V_k \cup \overline{V_k}$ and $V_k \cap \overline{V_k} = \emptyset$.

The edges crossing the cut are given by

1) $E \cap \{(s,i) | i \in \overline{V}_k\}$;
2) $E \cap \{(j,t) | j \in V_k\}$;
3) $E \cap \{(j,i) | j \in V_k \text{ and } i \in \overline{V_k}\}$.

In other words, the source node $s$ and the relay nodes $V_k$ are on one side of the cut, whereas the relay nodes $\overline{V_k}$ and the terminal node $t$ on the other side of the cut. The total capacity of an $s$-$t$-cut of size $k$ is given by

$$C_k = \sum_{i \in \overline{V_k}} C_{si} + \sum_{j \in V_k} \sum_{i \in \overline{V_k}} C_{ji} + \sum_{j \in V_k} C_{jt}. \tag{1}$$

## IV. Bounds and Results

In this section, we bound the network coding capacity of a connectivity graph, where the connections of the relay nodes form an instance of a quasi random geometric graph.

Let $G = (V, E)$ be a connectivity graph such that the vertex set $V$ consists of a source node $s$, a set of terminal nodes $T$, and a set of relay nodes $R$, that is, $V = \{s\} \cup T \cup R$. Recall that two nodes $u$ and $v$ in $G$ are connected by an edge with probability 1 if $d(u, v) \le r$, with probability $p$ if $r < d(u, v) \le r'$, and with probability 0 otherwise. Therefore, the probability $p'$ that two nodes $u$ and $v$ are connected can be bounded by

$$\frac{1}{4}\left(\pi r^2 + \pi(r'^2 - r^2)p\right) \le p' \le \pi r^2 + \pi(r'^2 - r^2)p. \tag{2}$$

The motivation for the lower bound stems from the fact that one of the nodes might be located in one of the corners of the unit square. The upper bound is a straightforward consequence of our connectivity rules.

These elementary observations allow us to bound the expected value of the cut $C_k$. By equation (1), we have

$$\begin{aligned} \mathbf{E}[C_k] &= \sum_{i \in \overline{V_k}} \mathbf{E}[C_{si}] + \sum_{j \in V_k} \sum_{i \in \overline{V_k}} \mathbf{E}[C_{ji}] + \sum_{j \in V_k} \mathbf{E}[C_{jt}] \\ &= p'(n + k(n-k)). \end{aligned}$$

In particular, $\mathbf{E}[C_k] = \mathbf{E}[C_{n-k}]$ holds for all $k$ in the range $0 \le k \le n$. Furthermore, we have

$$\mathbf{E}[C_0] = \mathbf{E}[C_n] \le \mathbf{E}[C_1] = \mathbf{E}[C_{n-1}] \le \cdots \le \mathbf{E}[C_{\lceil n/2 \rceil}].$$

Our goal is to prove that the capacity $C_k$ of an $s$-$t$-cut is concentrated around its expected value. A technical difficulty arises because the edges between relay nodes in the graph $G$ are in general not mutually independent. Indeed, if two relay nodes $u$ and $v$ are connected, and $u$ is connected to yet another relay node $w$, then there is a good chance that $v$ is connected to $w$. Put differently, we have

$$\Pr[(v,w) \in E | (u,v) \in E, (u,w) \in E] > \Pr[(v,w) \in E],$$

whence the three events $(u,v) \in E$, $(u,w) \in E$, and $(v,w) \in E$ are not independent. In Fig. 3, we sketch different geometric situations between two nodes; positioning a node $w$ within the transmission range of $u$ nicely illustrates the intuition behind this fact.

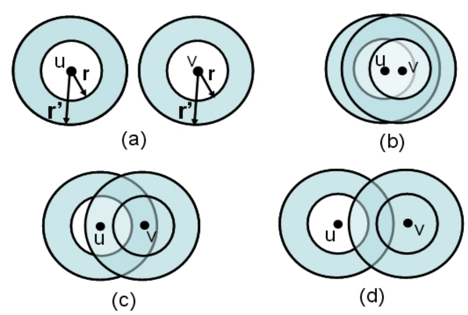


Fig. 3. Consider two relay nodes $u$ and $v$ of $G|_R$. Subfigure (a) illustrates the situation when the two nodes are not connected and far apart. The other subfigures illustrate the following situations: (b) $d(u,v) \le r$, (c) $r < d(u,v) \le r'$, and (d) $2r < d(u,v) \le r + r'$.

However, certain edges in a connectivity graph are independent. Indeed, all edges that are incident with a fixed (common) vertex are independent, since the coordinates of the vertices in the underlying quasi geometric random graph are chosen independently and uniformly at random. Consequently, the random variables in the set $\{C_{ij} \,|\, j \in I\}$, where $i$ is fixed, are independent. We will take advantage of this fact in our proof of the concentration result. To that end, recall Chernoff's bound for sums of independent Bernoulli random variables.

*Lemma 6 (Chernoff bound):* Let $X_1, \dots, X_m$ be independent Bernoulli random variables such that $\Pr[X_k = 1] = p'$ and $\Pr[X_k = 0] = 1 - p'$, and let $X = \sum_{k=1}^m X_k$. For $0 < \epsilon < 1$, we have

$$\Pr[X \leq (1-\epsilon)\mathbf{E}[X]] \leq \exp(-\mathbf{E}[X]\epsilon^2/2).$$

*Proof:* See, for instance, [13, p. 66] for a proof of this well-known bound. ∎

In the proof of the next result, we are going to take advantage of the following simple fact. Let $Z_1, \dots, Z_{k+1}$ be $k+1$ random variables. Then the event $Z_1 + \cdots + Z_{k+1} \leq \epsilon$ is contained in the event $\bigcup_{\ell=1}^{k+1} \big(Z_\ell \leq \epsilon/(k+1)\big)$; hence, the union bound yields

$$\Pr\left[\sum_{\ell=1}^{k+1} Z_\ell \leq \epsilon\right] \leq \sum_{\ell=1}^{k+1} \Pr\left[Z_\ell \leq \epsilon/(k+1)\right].$$

*Theorem 7:* The capacity $C_k$ of a cut of size $k$ in a quasi random geometric graph satisfies

$$\Pr[C_k \leq (1-\epsilon)\mathbf{E}[C_k]] \leq \exp\left(\ln(k+1) - \frac{\epsilon^2(n-k)p'}{8}\right).$$

*Proof:* Recall that the capacity $C_k$ of a cut

$$(\{s\} \cup V_k; \overline{V}_k \cup \{t\}),$$

is given by equation (1), which we can rewrite in the form

$$C_k = \sum_{i \in \overline{V}_k} C_{si} + \sum_{j \in V_k} \sum_{i \in \overline{V}_k \cup \{t\}} C_{ji}. \tag{3}$$

We split this sum into $k+1$ simpler parts. Let

$$Y_s = \sum_{i \in \overline{V}_k} C_{si} \quad \text{and} \quad Y_j = \sum_{i \in \overline{V}_k \cup \{t\}} C_{ji} \quad \text{for} \quad j \in V_k.$$

The random variable $Y_\ell$ counts the number of edges that are incident with the node $\ell \in \{s\} \cup V_k$. In a quasi random geometric graph, the edges incident with a fixed vertex $\ell$ are independent; hence, $Y_\ell$ is a sum of the independent Bernoulli random variables $C_{\ell i}$.

Equation (3) can now be written as a sum of the $k+1$ random variables $C_k = \sum_{\ell \in \{s\} \cup V_k} Y_\ell$. It follows that

$$\begin{aligned}
&\Pr[C_k - \mathbf{E}[C_k] \leq -\epsilon \mathbf{E}[C_k]] \\
&\leq \Pr\left[\bigcup_{\ell \in \{s\} \cup V_k} \left(Y_\ell - \mathbf{E}[Y_\ell] \leq -\frac{\epsilon}{k+1}\mathbf{E}[C_k]\right)\right] \\
&\leq \sum_{\ell \in \{s\} \cup V_k} \Pr\left[Y_\ell - \mathbf{E}[Y_\ell] \leq -\frac{\epsilon}{k+1}\mathbf{E}[C_k]\right]
\end{aligned}$$

Recall that $\mathbf{E}[Y_s] = p'(n-k)$ and $\mathbf{E}[Y_\ell] = p'(n-k+1)$, where $\ell \in V_k$. Using the Chernoff bound, we obtain

$$\begin{aligned}
&\Pr[Y_s - \mathbf{E}[Y_s] \leq -\epsilon \mathbf{E}[C_k]/(k+1)] \\
&\quad \leq \Pr[Y_s - \mathbf{E}[Y_s] \leq -\epsilon(k+1)\mathbf{E}[Y_s]/(k+1)] \\
&\quad \leq \exp(-\epsilon^2 \mathbf{E}[Y_s]/2) = \exp(-\epsilon(n-k)p'/2) \\
&\quad \leq \exp(-\epsilon(n-k)p'/8).
\end{aligned}$$

For a cut of size $k \geq 1$, we can obtain for the random variables $Y_\ell$ with $\ell \in V_k$ the following bound:

$$\begin{aligned}
&\Pr[Y_\ell - \mathbf{E}[Y_\ell] \leq -\epsilon \mathbf{E}[C_k]/(k+1)] \\
&\quad = \Pr[Y_\ell - \mathbf{E}[Y_\ell] \leq -\epsilon k \mathbf{E}[Y_\ell]/(k+1)] \\
&\quad \leq \exp(-\epsilon^2 \mathbf{E}[Y_\ell] k^2/(2(k+1)^2)) \\
&\quad \leq \exp(-\epsilon^2 (n-k)p'/8),
\end{aligned}$$

where the last inequality follows from the inequalities $1/4 \leq k^2/(k+1)^2$ for all $k \geq 1$, and $\mathbf{E}[Y_\ell] = p'(n-k+1) \geq p'(n-k)$.

The claim follows by combining these bounds. ∎

In the next two theorems, we are going to show that when the number $n$ of relay nodes is large, then the capacity of a minimum cut is—with high probability—concentrated about the value $np' = \mathbf{E}[C_0]$.

*Theorem 8:* Let $G$ be a connectivity graph with one source node $s$, $n$ relay nodes, and a set $T$ of terminal nodes. Then, with probability $1 - O(\tau/n^2)$, where $\tau = |T|$, the network coding capacity $C_{s,T}$ of $G$ is bounded from below by

$$C_{s,T} \geq (1-\epsilon)\mathbf{E}[C_0], \quad \text{where} \quad \epsilon = \sqrt{\frac{2^6 \ln n}{np'}},$$

where $p'$ satisfies (2).

*Proof:* Let $C_{min}(s,t)$ denote the capacity of a minimum $s$-$t$-cut. Let us assume further that this minimum cut has size $k$, that is, $C_{min}(s,t) = C_k$. By exchanging the role of $s$ and $t$ if necessary, we may assume that $k$ is in the range $0 \leq k \leq \lceil n/2 \rceil$. Since $\mathbf{E}[C_k] \geq \mathbf{E}[C_0]$ holds for all cut sizes $k$, we have

$$\begin{aligned}
\Pr[C_{min}(s,t) < (1-\epsilon)\mathbf{E}[C_0]] &\leq \Pr[C_k < (1-\epsilon)\mathbf{E}[C_k]] \\
&\leq \exp(\ln(k+1) - \epsilon^2(n-k)p'/8)
\end{aligned}$$

where the last inequality follows from the previous theorem. Substituting the value of $\epsilon$ from the hypothesis yields

$$\Pr[C_{min} < (1-\epsilon)\mathbf{E}[C_0]] = O(1/n^2).$$

Consequently, the probability that the network coding capacity $C_{s,T}$ will be below the value $(1-\epsilon)\mathbf{E}[C_0]$ can be bounded by

$$\begin{aligned}
&\Pr[C_{s,T} < (1-\epsilon)\mathbf{E}[C_0]] \\
&\qquad \leq \Pr\left[\bigcup_{t \in T} (C_{min}(s,t) < (1-\epsilon)\mathbf{E}[C_0])\right] \\
&\qquad \leq \sum_{t \in T} \Pr[C_{min}(s,t) < (1-\epsilon)\mathbf{E}[C_0]] \\
&\qquad = O(\tau/n^2),
\end{aligned}$$

as claimed. ∎

We complement the above lower bound by a high-probability upper bound on the network coding capacity.

*Theorem 9:* Let $G$ be a connectivity graph with one source node $s$, $n$ relay nodes, and a set $T$ of terminal nodes. Then, with probability $1 - O(1/n^{4/3})$, the network coding capacity $C_{s,T}$ of $G$ is bounded from above by

$$C_{s,T} \leq (1+\epsilon)\mathbf{E}[C_0], \quad \text{where} \quad \epsilon = \sqrt{\frac{4 \ln n}{\mathbf{E}[C_0]}}.$$

*Proof:* If the network coding capacity $C_{s,T}$ exceeds the value $(1+\epsilon)\mathbf{E}[C_0]$, then the capacity of any $s$-$t$-cut, for any $t \in T$, must exceed that value as well; in particular, the cut $(\{s\}; R \cup T)$ must have capacity exceeding $(1+\epsilon)\mathbf{E}[C_0]$. Since we assume that the source node is not directly connected to any terminal node, we obtain

$$\begin{aligned}&\Pr[C_{s,T} > (1+\epsilon)\mathbf{E}[C_0]] \\ &\quad \leq \Pr[\textstyle\sum_{r\in R} C_{sr} > (1+\epsilon)\mathbf{E}[C_0]] \\ &\quad \leq \Pr[|\textstyle\sum_{r\in R} C_{sr} - \mathbf{E}[C_0]| > \epsilon \mathbf{E}[C_0]]\end{aligned}$$

The indicator random variables $C_{sr}$, with $r \in R$, are mutually independent, as the location of the relay nodes are independently and identically distributed in the unit square. Recall that one version of the Chernoff bound for independent identically distributed indicator random variables $X_i$ with $\Pr[X_i = 1] = p'$ is given by $\Pr[|\sum_{i=1}^n X_i - np'| > t] < 2\exp(-t^2/3np')$. Applying this bound to the indicator random variables $C_{sr}$ yields

$$\begin{aligned}&\Pr[|\textstyle\sum_{r\in R} C_{sr} - \mathbf{E}[C_0]| > \epsilon\mathbf{E}[C_0]] \\ &\qquad < 2\exp(-\epsilon^2 \mathbf{E}[C_0]^2/(3\mathbf{E}[C_0])) \\ &\qquad = 2\exp\left(-\frac{4\ln n}{E[C_0]}\frac{\mathbf{E}[C_0]^2}{3\mathbf{E}[C_0]}\right) \\ &\qquad = O(n^{-4/3}),\end{aligned}$$

which proves the claim. ■

*Remark 10:* Our results easily generalize to more general substrates of unit area (not just unit squares), as long as the assumption holds that the nodes are uniformly distributed over the area. The concentration results are not affected by such a change, but the connectivity probability $p'$ might be dramatically different. For instance, if the area is a rectangle that is $\varepsilon$ high and $1/\varepsilon$ wide, then $p'$ approaches 0 as $\varepsilon$ approaches 0.

## V. Simulations and Experiments

We conducted simulations for various instances of $\mathcal{G}^{QRGG}$ using different parameters. Our simulation results support the high probability bounds on the network coding capacity given in Theorems 9 and 10.

In a first experiment, we determined the minimum capacity of an $s$-$t$ cut for different instances of a connectivity graph in $\mathcal{G}^{QRGG}$ with a fixed number of nodes. Fig. 4 shows the results of such an experiment with $n = 200$ relay nodes. The radio transmission range is chosen such that within a radius of $r = 0.1$ the connectivity is guaranteed and up to a radius of $r' = 0.2$ one might get connected. The plot shows that the capacity of the network is concentrated around the expected value of 13 which is in agreement with Theorem 9 and 10 for the above values of $n$, $r$ and $r'$.

Fig. 5 shows the result of a second experiment. This time, the number of relay nodes is once again $n = 200$, but the transmission range is higher, namely the inner radius equals $r = 0.13$ and outer radius equals $r' = 0.18$. We generated random instances of $\mathcal{G}^{QRGG}$ with these parameters and determined the minimum cut. One can easily see that the capacity of the network is likely to be higher, as expected.

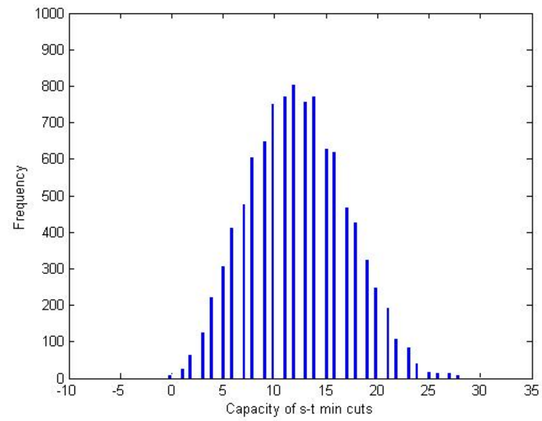


Fig. 4. n=200, r=0.1, $r' = 0.2$

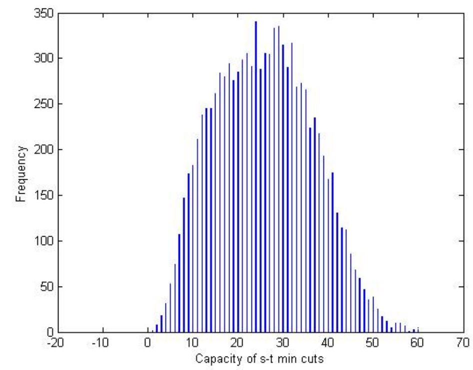


Fig. 5. n=200, r=0.13, $r' = 0.18$

For larger $n$, we could observe that the histograms become more concentrated around the expected capacity of a minimum cut, as predicted by our theory.

In a third series of experiments, we simulated the increase of capacity of the minimum cut for different values of $r$ and $n$. In this case, we also modeled the connectivity probability as a decreasing function of distance, following Remark 3,

$$p = \left(1 - \sqrt{\frac{d(i,j)^2 - r^2}{r'^2 - r^2}}\right) p_{connection},$$

where $d(i,j)$ is the Euclidean distance between any two nodes $i$ and $j$ such that $r < |d(i,j)| < r'$, and $p_{connection}$ is a probability that accounts for the interference noise in communication.

As it can be seen from Fig. 6, the value of the capacity grows more rapidly for lower values of $r$. This is intuitive because in that case not many nodes are connected for small values of $n$. As we increase $n$ but keep $r$ constant, the capacity of the minimum cut must increase, since more and more nodes are packed in the same area.

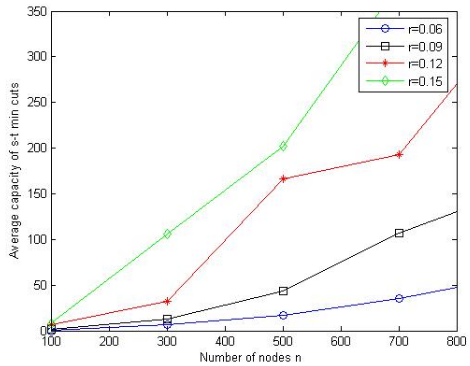


Fig. 6. The capacity of s-t minimum cuts with different values of $n$ and $r$

## VI. CONCLUSION

We modeled a wireless network using quasi random geometric graphs and showed that the capacity of the minimum cut of network coding is concentrated around the value $np' = \mathbf{E}[C_0]$, where $n$ denotes the number of relay nodes and $p'$ denotes the average connection probability. We obtained high probability bounds for this model without making simplifying assumption about the independence of the edges. More realistic models (for example, when the probability of connectivity drops exponentially with distance to account for signal attenuation) can be easily incorporated into our framework without changing the theory in a significant way.

## REFERENCES


[1] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung. Network information flow. *IEEE Trans. Info. Theory*, 46:1204–1216, 2000.

[2] N. Alon and J. H. Spencer. *The Probabilistic Method*. John Wiley, 2000.

[3] Reinhard Diestel. *Graph Theory*. Graduate Texts in Mathematics, Volume 173. Springer-Verlag, Heidelberg, third edition edition, July 2005.

[4] C. Fragouli, J. Le Boudec, and J. Widmer. Network coding: An instant primer. *ACM SIGCOMM Computer Communication Review*, 36(1):63–68, 2006.

[5] T. Ho, R. Koetter, M. Medard, D. Karger, and M. Effros. The benefits of coding over routing in a randomized setting. *In Proceedings of the IEEE International Symposium on Information Theory*, page 442, Yokohama, Japan, June 2003.

[6] T. Ho, M. Medard, R. Koetter, D. R. Karger, M. Effros, J. Shi, and B. Leong. Toward a random operation of networks. *submitted to IEEE Trans. Inform. Theory.*, 2006.

[7] T. Ho, M. Medard, J. Shi, M. Effros, and D. R. Karger. On randomized network coding. *41st Annual Allerton Conference on Communication Control and Computing*, Oct. 2003.

[8] S. Jaggi, P. Sanders, P. A. Chou, M. Effors, S. Egner, and L. M. Tolhuizen. Polynomial time algrithms for multicast network code construction. *IEEE Trans. Info. Theory*, 51(6), June 2005.

[9] D. R. Karger. Random sampling in cut, flow and network design problems. *Math. of Oper. Res.*, 24(2):0383 0413, 1999.

[10] R. Koetter and M. Medard. An algebraic approach to network coding. *IEEE Trans. on Networking*, October 2003.

[11] F. Kuhn, R. Wattenhofer, and A. Zollinger. Ad-hoc networks beyond unit disk graphs. *1st ACM Joint Workshop on Foundations of Mobile Computing (DIALM-POMC), San Diego, California, USA*, 2003.

[12] S.-Y. R. Li, R. W. Yeung, and N. Cai. Linear network coding. *IEEE Trans. Inform. Theory*, IT-49(2):371381, Feb. 2003.

[13] M. Mitzenmacher and E. Upfal. *Probability and Computing: Randomized Algorithms and Probabilistic Analysis*. Cambridge University Press, 2005.

[14] A. Ramamoorthy, J. Shi, and R.D. Wesel. On the capacity of network coding for random networks. In *Allerton Conference on Communication,Control and Computing*, 2003.

[15] A. Ramamoorthy, J. Shi, and R.D. Wesel. On the capacity of network coding for random networks. *IEEE Trans. Info. Theory*, 51(8), Aug. 2005.

[16] R. W. Yeung. *A First Course in Information Theory*. Kluwer/Plenum, 2002.