Capacity of Multi-Channel Wireless Networks with Random (c, f) Assignment

Technical Report (June 2007)

Vartika Bhandari
Dept. of Computer Science, and
Coordinated Science Laboratory
University of Illinois at Urbana-Champaign
vbhandar@uiuc.edu

Nitin H. Vaidya
Dept. of Electrical and Computer Eng., and
Coordinated Science Laboratory
University of Illinois at Urbana-Champaign
nhv@uiuc.edu

Abstract

The issue of transport capacity of a randomly deployed wireless network under random (c, f) channel assignment was considered by us in [1]. We showed in [1] that when the number of available channels is $c = O(\log n)$, and each node has a single interface assigned a random f subset of channels, the capacity is $\Omega(W\sqrt{\frac{f}{cn\log n}})$ and $O(W\sqrt{\frac{p_{rnd}}{n\log n}})$, and conjectured that capacity is $\Theta(W\sqrt{\frac{p_{rnd}}{n\log n}})$. We now present a lower bound construction that yields capacity $\Omega(W\sqrt{\frac{p_{rnd}}{n\log n}})$ for all $c = O(\log n)$ and $2 \leq f \leq c$. This establishes the capacity under random (c, f) assignment as $\Theta(W\sqrt{\frac{p_{rnd}}{n\log n}})$. The surprising implication of this result is that when $f = \Omega(\sqrt{c})$, random (c, f) assignment yields capacity of the same order as attainable via unconstrained switching. Also of interest is the routing/scheduling procedure we utilize to achieve capacity, which marks a significant point of departure from the construction used to obtain the previous lower bound of $\Omega(W\sqrt{\frac{f}{cn\log n}})$. This procedure requires synchronized route-construction for all flows in the network, leading to the open question of whether it is possible to achieve capacity using asynchronous routing/scheduling procedures.

I. Introduction

In [1], we argued for the need to study the performance of multi-channel networks in situations where there are constraints on channel switching. We proposed some constraint models in [1] to capture some expected constraints, and analyzed two such models, viz., adjacent (c, f) assignment and random (c, f) assignment. We studied the impact of such restricted switching, quantified by the parameter f (where f is the number of channels an individual node may switch to) in the regime $c = O(\log n)$. One of our proposed models was termed random (c, f) assignment. For this model, we proved in [1] that capacity is $O(W\sqrt{\frac{p_{rnd}}{n\log n}})$ and $\Omega(W\sqrt{\frac{f}{cn\log n}})$.

In this paper, we establish the per-flow capacity under random (c, f) assignment for the regime $c = O(\log n)$ as $\Theta(W\sqrt{\frac{p_{rnd}}{n\log n}})$ by presenting a lower bound construction that yields $\Omega(W\sqrt{\frac{p_{rnd}}{n\log n}})$ per-flow throughput. It can be

This research is supported in part by NSF grant CNS 06-27074, US Army Research Office grant W911NF-05-1-0246, and a Vodafone Graduate Fellowship.

This report is a revised version of an identically named technical report dated January 2007, and supercedes it.

shown that $p_{rnd} \geq 1 - e^{-\frac{f^2}{c}}$. Hence the implication of this result is that when $f = \Omega(\sqrt{c})$, random (c, f) assignment yields capacity of the same order as attainable via unconstrained switching. Thus, for the random (c, f) assignment model, $\sqrt{c}$-switchability is sufficient to make order-optimal use of all c channels, when $c = O(\log n)$.

Interestingly, our capacity achieving routing/scheduling procedure requires that routes/schedules for all flows be computed in lock-step in a synchronized manner. This leaves open the question of whether capacity can be achieved via asynchronous routing/scheduling procedures.

II. Notation and Terminology

Throughout this paper, we use the following standard asymptotic notation [2]:

- $f(n) = O(g(n))$ means that $\exists c, N_o,$ such that $f(n) \leq cg(n)$ for $n > N_o$
- $f(n) = o(g(n))$ means that $\lim_{n\to\infty} \frac{f(n)}{g(n)} = 0$
- $f(n) = \omega(g(n))$ means that $g(n) = o(f(n))$
- $f(n) = \Omega(g(n))$ means that $g(n) = O(f(n))$
- $f(n) = \Theta(g(n))$means that $\exists c_1, c_2, N_o,$ such that $c_1 g(n) \leq f(n) \leq c_2 g(n)$ for $n > N_o$

When $f(n) = O(g(n))$, any function $h(n) = O(f(n))$ is also $O(g(n))$. We often refer to such a situation as $h(n) = O(f(n)) \implies O(g(n))$.

As in [3], we say that the per flow network throughput is $\lambda(n)$ if each flow in the network can be guaranteed a throughput of at least $\lambda(n)$ with probability 1, as $n \to \infty$.

Whenever we use log without explicitly specifying the base, we imply the *natural* logarithm.

III. Some Useful Results

Theorem 1: (Vapnik-Chervonenkis Theorem) Let S be a set with finite VC dimension VCdim$(\mathcal{S})$. Let $\{X_i\}$ be i.i.d. random variables with distribution P. Then for $\varepsilon, \delta > 0$:

$$Pr\left(\sup_{D\in\mathcal{S}} \left|\frac{1}{N}\sum_{i=1}^{N} I_{X_i \in D} - P(D)\right| \leq \varepsilon\right) > 1 - \delta$$
$$\text{whenever } N > \max\left(\frac{8\text{VCdim}(\mathcal{S})}{\varepsilon}\log_2\frac{16e}{\varepsilon}, \frac{4}{\varepsilon}\log_2\frac{2}{\delta}\right)$$

Theorem 2: (Chernoff Bound [4]) Let $X_1, \ldots, X_n$ be independent Poisson trials, where $Pr[X_i = 1] = p_i$. Let $X = \sum_{i=1}^{n} X_i$. Then, for any $\beta > 0$:

$$Pr[X \geq (1+\beta)E[X]] < \left(\frac{e^\beta}{(1+\beta)^{(1+\beta)}}\right)^{E[X]} \tag{1}$$

Theorem 3: (Chernoff Upper Tail Bound [4]) Let $X_1,\ldots,X_n$ be independent Poisson trials, where $Pr[X_i = 1] = p_i$. Let $X = \sum_{i=1}^{n} X_i$. Then, for $0 < \beta \leq 1$:

$$Pr[X \geq (1+\beta)E[X]] \leq \exp(-\frac{\beta^2}{3}E[X]) \tag{2}$$

Theorem 4: (Chernoff Lower Tail Bound [4]) Let $X_1,\ldots,X_n$ be independent Poisson trials, where $Pr[X_i = 1] = p_i$. Let $X = \sum_{i=1}^{n} X_i$. Then, for $0 < \beta < 1$:

$$Pr[X \leq (1-\beta)E[X]] \leq \exp(-\frac{\beta^2}{2}E[X]) \tag{3}$$

Lemma 1: The chernoff bounds continue to apply if the Poisson trials are not independent, but are negatively correlated.

This is a well-known, and often-used result, e.g., see [5].

Lemma 2: Suppose we are given a unit toroidal region with n nodes located uniformly at random, and the region is sub-divided into axis-parallel square cells of area $a(n)$ each. If $a(n) = \frac{100\alpha(n)\log n}{n}, 1 \leq \alpha(n) \leq \frac{n}{100\log n}$, then each cell has at least $(100\alpha(n) - 50)\log n$, and at most $(100\alpha(n) + 50)\log n$ nodes, with high probability.

Proof: It is known that the set of axis-parallel squares in $\mathbb{R}^2$ has VC-dimension 3. In our construction, we have a set of axis-parallel square cells $\mathcal{S}$ such that the cells all have area $a(n) = \frac{100\alpha\log n}{n}$. Then considering the n random variables X_i denoting node positions, $Pr[X_i \in D(D \in \mathcal{S})] = \frac{100\alpha\log n}{n}$. Then, from the VC-theorem (Theorem 1):

$$Pr\left(\sup_{D\in\mathcal{S}}\left|\frac{\text{No. of nodes in } D}{n} - \frac{100\alpha(n)\log n}{n}\right| \leq \varepsilon(n)\right) > 1 - \delta(n)$$
$$\text{whenever } n > \max\left(\frac{24}{\varepsilon}\log_2\frac{16e}{\varepsilon}, \frac{4}{\varepsilon}\log_2\frac{2}{\delta}\right)$$

This is satisfied when $\varepsilon(n) = \delta(n) = \frac{50\log n}{n}$. Thus, with probability at least $1 - \frac{50\log n}{n}$, the population $Pop(D)$ of cell D satisfies:

$$\frac{(100\alpha(n) - 50)\log n}{n} \leq \text{Pop(D)} \leq \frac{(100\alpha(n) + 50)\log n}{n} \tag{4}$$

■

Lemma 3: Suppose we are given a unit toroidal region with n points(or nodes) located uniformly at random, let us consider the set of all circles of radius R and area $A(n) = \pi R^2$ on the unit toroid. If $A(n) = \frac{100\alpha(n)\log n}{n}, 1 \leq \alpha(n) \leq \frac{n}{100\log n}$, then each circle has at least $(100\alpha(n) - 50)\log n$, and at most $(100\alpha(n) + 50)\log n$ of these points (or nodes), with high probability.

Proof: The set of all circles of radius R in $\mathbb{R}^2$ has VC-dimension 3 (e.g., see [3]). Thereafter by the same argument as in the proof of Lemma 2, the result proceeds. ■

Lemma 4: If n pairs of points (P_i, Q_i) are chosen uniformly at random in the unit area network, the resultant set of straight-line formed by each pair $L_i = P_iQ_i$ satisfies the condition that no cell has more than $n\sqrt{a(n)}$ lines passing through it w.h.p.

Proof: Given the lines L_i are i.i.d., the proof of Lemma 3 in [6] can be applied to prove this result. ∎

Lemma 5: For all $0 \leq x \leq 1 : (1-x) \leq e^{-x}$.

Theorem 5: (Hall's Marriage Theorem [7], [8]) Given a set $\mathcal{S}$, let $\mathcal{T} = \{\mathcal{T}_1, \mathcal{T}_2, \dots \mathcal{T}_n\}$ be a finite system of subsets of $\mathcal{S}$. Then $\mathcal{T}$ possesses a system of distinct representatives if and only if for each k in $1,2,..,n$, any selection of k of the sets $\mathcal{T}_i$ will contain between them at least k elements of $\mathcal{S}$. Alternatively stated: for all $\mathcal{A} \subseteq \mathcal{T}$, the following is true: $|\cup\mathcal{A}| \geq |\mathcal{A}|$

Lemma 6: The number of subsets of size k chosen from a set of m elements is given by $\binom{m}{k} \leq \left(\frac{me}{k}\right)^k$.

Theorem 6: (Integrality Theorem [2]) If the capacity function of a network flow graph takes on only integral values (i.e., each edge has integer capacity), then the maximum flow x produced by the Ford-Fulkerson method has the property that $|x|$ is integer-valued. Moreover, for all vertices u and v, the value of $x(u,v)$ is an integer.

IV. Network Model

We consider a network of n *single-interface* nodes randomly deployed over a unit torus. Each node is the source of exactly one flow. As in [3], each source S selects a destination by first fixing on a point D' uniformly at random, and then picking the node D (other than itself), that is closest to D'. The total bandwidth (data-rate) available is W, and it is divided into c channels of equal bandwidth $\frac{W}{c}$, where $c = O(\log n)$. We assume that $c \geq 2$, as $c = 1$ implies that $f = 1$ is the only possibility, which yields the degenerate single-channel case. We also assume $2 \leq f \leq c$. A justification for not allowing $f = 1$ for $c \geq 2$ is given in [1], [9], where we show that for the random (c, f) model (and also the adjacent (c, f) model described in [1]), $f = 1$ and $c \geq 2$ leads to zero capacity, as some flow will get no throughput w.h.p.

V. Some results about the traffic model

In this section we establish that for the traffic model of [3] (which is also used in this paper), a node is the destination of $O(\log n)$ flows w.h.p. Also, at least one node is the destination of $\Omega(\log n)$ flows w.h.p.

Lemma 7: The number of flows for which any node is the destination is $O(\log n)$ w.h.p.

Proof: Consider a flow's pseudo-destination D'. Consider a circle of radius $\sqrt{\frac{100\log n}{\pi n}}$, and hence area $\frac{100\log n}{n}$ centered around this pseudo-destination. From Lemma 3, all such circles contain $\Theta(\log n)$ nodes, w.h.p. In a rare scenario, one of these nodes could potentially be the source node for that flow. However, the circle still has more than one node other than the flow's source. Thus, the flow will select as its destination, some node within this circle. Hence a flow can only be assigned a destination within distance $\sqrt{\frac{100\log n}{\pi n}}$ from its pseudo-destination. Thus it proceeds that a node can only be the destination for flows whose pseudo-destination lies within a distance $\sqrt{\frac{100\log n}{\pi n}}$ from it. From Lemma 3, each circle of this size contains $O(\log n)$ pseudo-destinations w.h.p. Thus no node is the destination of more than $O(\log n$ flows. ∎

Lemma 8: For large n, at least one node is a destination for $\Omega(\log n)$ flows with a probability at least $\frac{1}{e}(1-\frac{1}{e})(1-\delta)$, where $\delta > 0$ is an arbitrarily small constant.

Proof: The necessary condition for connectivity in [10] (Theorem 2.1 of [10]) is established by proving that if we consider $R(n)$ such that $\pi R^2(n) = \frac{\log n + b(n)}{n}$, where $\limsup b(n) = b < \infty$, then with positive probability, there exists at least one node x which is isolated, i.e., there is no other node within distance $R(n)$ of x. In the context of [10], this was utilized by interpreting $R(n)$ as transmission range, and thus obtaining a lower bound for connectivity. However, we now exploit that result in a different manner to prove our lemma as follows: Choose $R(n) = \sqrt{\frac{\log n+1}{\pi n}}$, i.e., $b(n) = b = 1$. Note that in this proof, $R(n)$ is *not* the transmission range; it is merely a chosen distance value. Then by invoking Theorem 2.1 from [10], there exists a node A such that there is no other node within a distance $R(n)$ from it, with probability p where $\liminf p \geq e^{-b}(1-e^{-b}) = \frac{1}{e}(1-\frac{1}{e})$. In fact, from the proof of Theorem 2.1 in [10], it proceeds that $p \geq (1-\varepsilon)\frac{1}{e}(1-\frac{1}{e})$, for any $\varepsilon > 0$, and sufficiently large n. Call this event $\mathcal{E}_1$.

Thus, given event $\mathcal{E}_1$ has occurred and such a node A exists, if we consider the Voronoi tessellation generated by the n nodes, then the Voronoi polygon of A has area at least $\pi(\frac{R(n)}{2})^2 = \frac{\pi R^2(n)}{4} = \frac{\log n+1}{4n}$. Note that this tessellation constitutes a spatial partition of the network area. Also, it immediately proceeds from the traffic model, that if a flow's pseudo-destination falls within the polygon of node x, then x is selected as that flow's destination, unless x is itself the source of that flow (since a generator (node) is always the nearest generator to points within its own polygon). Also recall that pseudo-destinations are chosen uniformly at random. Let $X_i, 1 \leq i \leq n$ be indicator variables such that $X_i = 1$ if A is flow i's destination, and 0 else. Then $Pr[X_i = 1] = 0$ if A is the source of flow i (and there is exactly one such i). For all other values of i, $Pr[X_i = 1|\mathcal{E}_1] \geq \frac{\log n+1}{4n}$, since A is selected as flow i's destination if either (1) flow i's pseudo-destination falls in A's Voronoi polygon (the probability of this event is given by the area of A's Voronoi polygon, and is thus at least $\frac{\log n+1}{4n}$, or (2) if flow i's pseudo-destination falls within the polygon of its own source, and A is the next-nearest node (we ignore this probability, as we only require a lower bound). Let $X = \sum X_i$. Thus $E[X|\mathcal{E}_1] \geq (1-\frac{1}{n})\frac{\log n+1}{4} \geq \frac{\log n}{4}$ for large n. The X_i's are i.i.d., and thus application of the Chernoff bound from Theorem 4, with $\beta = \frac{1}{2}$ yields that:

$$Pr[X \leq \frac{\log n}{8}|\mathcal{E}_1] \leq Pr[X \leq \frac{E[X]}{2}|\mathcal{E}_1] \leq \exp(-\frac{E[X]}{8}) \leq \exp(-\frac{\log n}{32}) = \frac{1}{n^{\frac{1}{32}}} \tag{5}$$

Denote by $\mathcal{E}_2$ the event that some node indeed is destination to at least $\frac{\log n}{8}$ flows. Then it proceeds from Eqn. (5) that $Pr[\mathcal{E}_2|\mathcal{E}_1] \geq 1 - \frac{1}{n^{\frac{1}{32}}}$. Also, $Pr[\mathcal{E}_2] \geq Pr[\mathcal{E}_1]Pr[\mathcal{E}_2|\mathcal{E}_1]$. Hence at least one node is a destination for $\Omega(\log n)$ flows with a probability at least $(1-\varepsilon)e^{-b}(1-e^{-b})(1-\frac{1}{n^{\frac{1}{32}}}) \geq \frac{1}{e}(1-\frac{1}{e})(1-\delta)$ for any chosen $\delta > \varepsilon$, and sufficiently large n. ■

VI. RANDOM (c, f) ASSIGNMENT

In this section we briefly describe the random (c, f) assignment model first described in [9], and summarize some already proven results that will be useful in proving the lower bound on capacity. In this assignment model, a node is assigned a subset of f channels uniformly at random from the set of all possible channel subsets of size f. Thus the probability that two nodes share at least one channel is given by $p_{rnd} = 1 - (1-\frac{f}{c})(1-\frac{f}{c-1})...(1-\frac{f}{c-f+1})$.

Lemma 9: For $c \geq 2$, and $2 \leq f \leq c$, the following holds:

$$\frac{cp_{rnd}}{f} \leq \min\{\frac{c}{f}, 2f\} \tag{6}$$

Proof: Since $p_{rnd} \leq 1$, we obtain that $\frac{cp_{rnd}}{f} \leq \frac{c}{f}$.

If $f \geq \sqrt{\frac{c}{2}}$, then $\frac{cp_{rnd}}{f} \leq \sqrt{2c} \leq 2f$ follows from the observation that $p_{rnd} \leq 1$. Hence, we focus on the case $f < \sqrt{\frac{c}{2}}$.

$$\begin{aligned} 1 - p_{rnd} &= (1-\frac{f}{c})(1-\frac{f}{c-1})...(1-\frac{f}{c-f+1}) \\ &\geq (1-\frac{f}{c-f+1})^f > (1-\frac{2f}{c})^f \geq 1-\frac{2f^2}{c} \\ &\therefore p_{rnd} \leq \frac{2f^2}{c} \\ &\therefore \frac{cp_{rnd}}{f} \leq 2f \end{aligned} \tag{7}$$

Thus, $\frac{cp_{rnd}}{f} \leq \min\{\frac{c}{f}, 2f\}$. ■

Lemma 10: $\min\{\frac{c}{f}, 2f\} \leq \sqrt{2c}$

Proof: For a given c, we have $2 \leq f \leq c$. Thus, given c, $\frac{c}{f}$ is a monotonically decreasing function of f, while $2f$ is a monotonically increasing function of f. $\frac{c}{f} = 2f = \sqrt{2c}$ at $f = \sqrt{\frac{c}{2}}$. For $f \leq \sqrt{\frac{c}{2}}$, $\min\{\frac{c}{f}, 2f\} = 2f \leq \sqrt{2c}$, and for $f > \sqrt{\frac{c}{2}}$, $\min\{\frac{c}{f}, 2f\} = \frac{c}{f} \leq \sqrt{2c}$. Thus $\min\{\frac{c}{f}, 2f\} \leq \sqrt{2c}$. ■

A. Sufficient Condition for Connectivity

This theorem has been stated and proved by us in [1] (also [9]). However, we repeat it here in the interests of clarity.

Theorem 7: With random (c, f) assignment, when $c = O(\log n)$, if $\pi r^2(n) = \frac{800\pi \log n}{p_{rnd} n}$, then:

$$Pr[\text{ network is connected }] \rightarrow 1$$

Proof: The construction is based on a notion of per-node *backbones*. Consider a subdivision of the toroidal unit area into square cells of area $a(n) = \frac{100 \log n}{p_{rnd} n}$. Then by setting $\alpha(n) = \frac{1}{p_{rnd}}$ in Lemma 2 there are at least $\frac{50 \log n}{p_{rnd}}$ nodes in each cell with probability at least $1 - \frac{50 \log n}{n}$. Set $r(n) = \sqrt{8a(n)}$. Then a node in any given cell has all nodes in adjacent cells within its range. Within each cell, choose $\frac{2 \log n}{p_{rnd}}$ nodes uniformly at random, and set them apart as *transition facilitators* (the meaning of this term shall become clear later). This leaves at least $\frac{48 \log n}{p_{rnd}}$ nodes in each cell that can act as *backbone candidates*.

Consider any node in any given cell. The probability that it can communicate to any other random node in its range is p_{rnd}. Then the probability that in an adjacent cell, there is no backbone candidate node with which it can communicate is less than $(1 - p_{rnd})^{\frac{48 \log n}{p_{rnd}}} \leq \frac{1}{e^{48 \log n}} = \frac{1}{n^{48}}$ (applying Lemma 5).

The probability that a given node cannot communicate with any node in some adjacent cell is thus at most $\frac{8}{n^{48}}$ (as there are upto 8 adjacent cells per node). By applying the union bound over all n nodes, the probability that at least one node is unable to communicate with any backbone candidate node in at least one of its adjacent cells is at most $\frac{8}{n^{47}}$.

We associate with each node x a set of nodes $\mathcal{B}(x)$ called the primary backbone for x. $\mathcal{B}(x)$ is constituted as follows. Throughout the procedure, cells that are already covered by the under-construction backbone are referred to as *filled* cells. x is by default a member of $\mathcal{B}(x)$, and its cell is the first *filled* cell. From each adjacent cell, amongst all backbone candidate nodes sharing at least one common channel with x, one node is chosen uniformly at random and added to $\mathcal{B}(x)$. Thereafter, from each cell bordering a filled cell, of all nodes sharing at least one common channel with some node already in $\mathcal{B}(x)$, one is chosen uniformly at random, and is added to $\mathcal{B}(x)$; the cell containing the chosen node gets added to the set of filled cells. This process continues iteratively, till there is one node from every cell in $\mathcal{B}(x)$. From our earlier observations, $\mathcal{B}(x)$ eventually covers all cells with probability at least $1-\frac{8}{n^{47}}$. Now consider any pair of nodes x and y. If $\mathcal{B}(x)\cap\mathcal{B}(y)\neq\phi$ the two nodes are obviously connected, as one can proceed from x on $\mathcal{B}(x)$ towards one of the intersection nodes, and thence to y on $\mathcal{B}(y)$, and vice-versa. Suppose, the two backbones are disjoint. Then x and y are still connected if there is some cell such that the member of $\mathcal{B}(x)$ in that cell (let us call it q_x) can communicate with the member of $\mathcal{B}(y)$ in that cell (let us call it q_y), either directly, or through a third node. q_x and q_y can communicate directly with probability 1 if they share a common channel. Thus the case of interest is one in which no cell has q_x and q_y sharing a channel.

If they do not share a common channel, we consider the event that there exists a third node amongst the *transition facilitators* in the cell through whom they can communicate. Note that, for two given backbones $\mathcal{B}(x)$ and $\mathcal{B}(y)$, the probability that in a network cell, given q_x and q_y that do not share a channel, they can both communicate with a third node z that did not participate in backbone formation and is known to lie in the same cell, is independent across cells. Therefore, the overall probability can be lower-bounded by obtaining for one cell the probability of q_x and q_y communicating via a third node z, given they have no common channel, considering that each cell has at least $\frac{2\log n}{p_{rnd}}$ possibilities for z, and treating it as independent across cells. We elaborate this further.

Let q_x have the set of channels $C(q_x)=\{c_{x_1},...,c_{x_f}\}$, and q_y have the set of channels $C(q_y)=\{c_{y_1},...,c_{y_f}\}$, such that $C(q_x)\cap C(q_y)=\phi$. Consider a third node z amongst the transition facilitators in the same cell as q_x and q_y. We desire z to have at least one channel common with both $C(q_x)$ and $C(q_y)$. Then let us merely consider the possibility that z enumerates its f channels in some order, and then inspects the first two channels, checking the first one for membership in $C(q_x)$, and checking the second one for membership in $C(q_y)$. This probability is $\left(\frac{f}{c}\right)\left(\frac{f}{c-1}\right)>\frac{f^2}{c^2}$. Thus q_x and q_y can communicate through z with probability $p_z>\frac{f^2}{c^2}=\Omega(\frac{1}{\log^2 n})$. [1] There are $\frac{2\log n}{p_{rnd}}$ possibilities for z within that cell, and all the possible z nodes have i.i.d. channel assignments. Thus, the probability that q_x and q_y cannot communicate through any z in the cell is at most $(1-p_z)^{\frac{2\log n}{p_{rnd}}}$, and the probability they can indeed do so is $p_{xy}\geq 1-(1-p_z)^{\frac{2\log n}{p_{rnd}}}$.

Thus, the probability that this happens in none of the $\frac{1}{a(n)}=\frac{p_{rnd}n}{100\log n}$ cells is at most $(1-p_{xy})^{\frac{p_{rnd}n}{100\log n}}\leq(1-p_z)^{\frac{2\log n}{p_{rnd}}\frac{p_{rnd}n}{100\log n}}<(1-\frac{f^2}{c^2})^{\frac{2\log n}{p_{rnd}}\frac{p_{rnd}n}{100\log n}}\leq e^{-\frac{f^2n}{50c^2}}\rightarrow e^{-\Omega(\frac{n}{\log^2 n})}$ (recall that $c=O(\log n)$). Applying union bound over all $\binom{n}{2}<\frac{n^2}{2}$ node pairs, the probability that some pair of nodes are not connected is at most $\frac{n^2e^{-\Omega(\frac{n}{\log^2 n})}}{2}<\frac{1}{2}e^{-\Omega(\frac{n}{\log^2 n})+2\log n}\rightarrow 0$. Applying union bound over this probability and the probability that some of the cells are not sufficiently populated (as mentioned earlier, this probability is at most $\frac{50\log n}{n}$), we obtain that the probability of a connected network converges to 1. ■

VII. Lower Bound on Capacity

We proved a lower bound of $\Omega(W\sqrt{\frac{f}{cn\log n}})$ for random (c,f) assignment in [1], [9]. From Lemma 9, it follows that $\frac{\sqrt{\frac{f}{cn\log n}}}{\sqrt{\frac{p_{rnd}}{n\log n}}}=\Omega(\frac{1}{\sqrt{f}})$. Thus for $f<100$, $\frac{\sqrt{\frac{f}{cn\log n}}}{\sqrt{\frac{p_{rnd}}{n\log n}}}=\Omega(1)$, and the construction presented in [1] (details in [9]) is

[1] As can be seen, this is a very loose bound on p_z and can be substantially improved. However, even this loose lower bound suffices for our current purpose.

asymptotically optimal.

We now present a construction that achieves $\Omega(W\sqrt{\frac{p_{rnd}}{n\log n}})$ when $f \geq 100$ (thus necessarily $c \geq 100$).

Subdivision of network region into cells: We use a square cell construction (similar to that used in [6], and subsequently in [11], [1]). The surface of the unit torus is divided into square cells of area $a(n)$ each, and the transmission range is set to $\sqrt{8a(n)}$, thereby ensuring that any node in a given cell is within range of any other node in any adjoining cell. Since we utilize the *Protocol Model* [3], a node C can potentially interfere with an ongoing transmission from node A to node B, only if $BC \leq (1+\Delta)r(n)$. Thus, a transmission in a given cell can only be affected by transmissions in other cells within a distance $(2+\Delta)r(n)$ from some point in that cell. Since Δ is independent of n, the number of cells that interfere with a given cell is only some constant (say β).

We choose $a(n) = \frac{250\max\{\log n, c\}}{p_{rnd} n} = \Theta(\frac{\log n}{p_{rnd} n})$ (since $c = O(\log n)$).

Then the following holds:

Lemma 11: Each cell has at least $\frac{4na(n)}{5} = \frac{200\max\{\log n, c\}}{p_{rnd}}$ and at most $\frac{6na(n)}{5} = \frac{300\max\{\log n, c\}}{p_{rnd}}$ nodes w.h.p.

Proof: We have chosen $a(n) = \frac{250\max\{\log n, c\}}{p_{rnd} n}$. Thus $a(n) \geq \frac{100\log n}{p_{rnd} n}$. Then if $c \leq \log n$, we can set $\alpha = \frac{2.5}{p_{rnd}} > 1$ in Lemma 2, and when $c > \log n$, i.e., $c = \gamma \log n (\gamma > 1)$ (recall that $c = O(\log n)$), we can set $\alpha = \frac{2.5\gamma}{p_{rnd}} > 1$, to obtain that the following holds with probability at least $1 - \frac{50\log n}{n}$ for all cells $\mathcal{D}$:

$$\frac{250\max\{\log n, c\}}{p_{rnd}} - 50\log n \leq Pop(\mathcal{D}) \leq \frac{250\max\{\log n, c\}}{p_{rnd}} + 50\log n$$

Thereafter noting that $\frac{250\max\{\log n, c\}}{p_{rnd}} - 50\log n \geq \frac{200\max\{\log n, c\}}{p_{rnd}}$, and $\frac{250\max\{\log n, c\}}{p_{rnd}} + 50\log n \leq \frac{300\max\{\log n, c\}}{p_{rnd}}$, completes the proof. ■

Corollary 1: Each cell has at least $\frac{200\log n}{p_{rnd}}$ nodes w.h.p.

Many of the intermediate results in the rest of this paper assume that the high-probability event of Lemma 11 holds.

We also state the following facts:

$$\frac{f}{c} \leq p_{rnd} \leq 1 \tag{8}$$

For large n, since $c = O(\log n)$, and $2 \leq f \leq c$:

$$\begin{aligned} na(n) &= \frac{250\max\{\log n, c\}}{p_{rnd}} = O(\log^2 n) \\ \frac{n\sqrt{a(n)}}{c} &= \frac{1}{c}\sqrt{\frac{250n\max\{\log n, c\}}{p_{rnd}}} = \Omega(\sqrt{\frac{n}{\log n}}) \\ \therefore f(n) &= O(na(n)) \implies f(n) = O(\frac{n\sqrt{a(n)}}{c}) \end{aligned} \tag{9}$$

$$\frac{1}{\sqrt{a(n)}} = \sqrt{\frac{p_{rnd}n}{250\max\{\log n, c\}}} = O(\sqrt{\frac{n}{\log n}})$$
$$\frac{n\sqrt{a(n)}}{c} = \frac{1}{c}\sqrt{\frac{250n\max\{\log n, c\}}{p_{rnd}}} = \Omega(\sqrt{\frac{n}{\log n}}) \tag{10}$$
$$\therefore f(n) = O(\frac{1}{\sqrt{a(n)}}) \implies f(n) = O(\frac{n\sqrt{a(n)}}{c})$$

Some properties of $SD'D$ routing: Recall that we use the traffic model of [3], where each source S first chooses a pseudo-destination D', and then selects the node D nearest to it as the actual destination. In [3], the route $SD'D$ was followed, whereby the flow traversed cells intersected by the straight line SD', and then took an extra last hop if required. In our case, it may not always suffice to use $SD'D$ routing (we elaborate on this later). However, this is still an important component of our routing procedure, and so we state and prove the following lemmas (some were also stated by us in [9]) for $SD'D$ routing:

Lemma 12: Given only straight-line SD' routing (no additional last-hop), the number of flows that enter any cell on their i-th hop is at most $\lfloor \frac{5na(n)}{4} \rfloor$ w.h.p., for any i.

Proof: Let us consider the straight-line part SD' of an $SD'D$ route. Thus all the n SD' lines are i.i.d. Denote by X_i^k the indicator variable which is 1 if the flow k enters a cell $\mathcal{D}$ on its i-th hop. Then, as observed in [6] (proof of Lemma 3), for i.i.d. straight lines, the X_i^k's are identically distributed, and X_i^k and X_j^l are independent for $k \neq l$. However for a given flow k, at most one of the X_i^k's can be 1 as a flow only traverses a cell once. Then $Pr[X_i^k = 1] = a(n) = \frac{250\max\{\log n, c\}}{p_{rnd}n}$.

Let $X_i = \sum_{k=1}^{n} X_i^k$. Then $E[X_i] = na(n)$. Also, for a given i, the X_i^k's are independent [6]. Then by application of the Chernoff bound from Thereom 3 (with $\beta = \frac{1}{4}$):

$$Pr[X_i \geq \frac{5E[X_i]}{4}] \leq \exp(-\frac{E[X_i]}{48})$$
$$\therefore Pr[X_i \geq \frac{1250\max\{\log n, c\}}{4p_{rnd}}] \tag{11}$$
$$\leq \exp(-\frac{250\max\{\log n, c\}}{48p_{rnd}}) < \frac{1}{n^5}$$

The maximum value that i can take is $\frac{2}{\sqrt{a(n)}} = \sqrt{\frac{2np_{rnd}}{250\max\{\log n, c\}}} < n$. Also the number of cells is $\frac{1}{a(n)} \leq n$. Then by application of union bound over all i, and all cells $\mathcal{D}$, the probability that $X_i \geq \frac{5E[X_i]}{4}$ is less than $\frac{1}{n^3}$, and thus the number of flows that enter any cell on any hop is less than $\frac{5na(n)}{4} = \frac{1250\max\{\log n, c\}}{4p_{rnd}}$ with probability at least $1 - \frac{1}{n^3}$. Resultantly, since X_i is an integer, we can say that it is at most $\lfloor \frac{5na(n)}{4} \rfloor$ w.h.p. ■

Lemma 13: The number of flows for which any single node is the destination is $O(na(n))$ w.h.p.

Proof: From Lemma 7, the number of flows for which any node is the destination is $O(\log n)$. We have chosen $a(n) = \Theta(\frac{\log n}{p_{rnd}n}) = \Omega(\frac{\log n}{n})$. Thus, $O(\log n) \implies O(na(n))$. This yields the result. ■

Lemma 14: If a node is destination of some flow, that flow's pseudo-destination must lie within either the same cell, or an adjacent cell w.h.p.

Proof: It was shown in the proof of Lemma 7 that a flow will be assigned to a destination lying within a circle of radius $\sqrt{\frac{100\log n}{\pi n}}$ centered around the pseudo-destination w.h.p. Conversely, if a flow is assigned to a node, then the pseudo-destination must lie within a circle of of radius $\sqrt{\frac{100\log n}{\pi n}}$ centered around the node.

It is easy to see that a circle of radius $\sqrt{\frac{100\log n}{\pi n}}$ centered at a node will fall completely within the cells adjacent to the node's cell (by our choice of cell-area $a(n)$). Hence if a node is destination of some flow, that flow's pseudo-destination must lie within either the same cell, or an adjacent cell. ■

Lemma 15: The number of $SD'D$ routes that traverse any cell is $O(n\sqrt{a(n)})$ w.h.p.

Proof: The proof for this lemma is largely based on a proof in [6]. Consider a cell $\mathcal{D}$. From Lemma 4 (which proceeds from a lemma in [6]) we know that the number of SD' straight-lines traversing any single cell are $O(n\sqrt{a(n)})$. We must now consider the number of routes whose last $D'D$ hop may enter this cell $\mathcal{D}$. If D is in the same cell as D', there is no extra hop. Let us now consider the case that D' lies in one of the 8 adjacent cells, but D lies in the cell $\mathcal{D}$ (from Lemma 14, we know that D lies in cell $\mathcal{D}$ only if D' lies in $\mathcal{D}$ or its adjacent cells). The number of flows for which D' lies in one of the 8 cells adjacent to $\mathcal{D}$ is $O(na(n))$ w.h.p.(by applying Lemma 2 to the set of n pseudo-destinations). Also from Eqn. (9), and the fact that $c > 1$, we know that $O(na(n)) \implies O(n\sqrt{a(n)})$. Thus the total number of traversing routes is $O(n\sqrt{a(n)})$. ■

Having stated and proved these lemmas, we now establish some properties of the spatial distribution of channels, and thereafter describe our scheduling/routing procedure further:

Definition 1: We define a term M_u where $M_u = \lceil \frac{9fna(n)}{25c} \rceil = \lceil \frac{90f\max\{\log n,c\}}{cp_{rnd}} \rceil$.

Then the following holds:

Lemma 16: If there are at least $\frac{200\max\{\log n,c\}}{p_{rnd}}$ nodes in every cell, of which we choose $\frac{180\max\{\log n,c\}}{p_{rnd}}$ nodes uniformly at random as *candidates* to examine, then, in each cell, amongst those $\frac{180\max\{\log n,c\}}{p_{rnd}}$ *candidate* nodes, at least $c - \lfloor \frac{f}{4} \rfloor$ channels have at least M_u nodes capable of switching on them, w.h.p.

Proof: Consider any single cell D. Let us denote by $\mathcal{E}$ the set of $\frac{180\max\{\log n,c\}}{p_{rnd}}$ nodes lying in cell D that are chosen uniformly at random for examination. Denote by I_{ji} the indicator variable that is 1 if a node j can switch on channel i and 0 else. $Pr[I_{ji} = 1] = \frac{f}{c}$ and $X_i = \sum_{j\in\mathcal{E}} I_{ji}$ is the number of nodes in $\mathcal{E}$ capable of switching on channel i. Then $E[X_i] = \frac{f}{c}\frac{180\max\{\log n,c\}}{p_{rnd}}$, and we can see that $M_u = \lceil \frac{E[X_i]}{2} \rceil$.

In light of Lemma 9, this leads to the following equations:

$$E[X_i] = \frac{180f\max\{\log n,c\}}{cp_{rnd}} \tag{12}$$

$$E[X_i] \geq \frac{180\max\{\log n,c\}}{\min\{2f,\frac{c}{f}\}} \geq \frac{90\max\{\log n,c\}}{f} \tag{13}$$

$$E[X_i] \geq 180f \text{ from Eqn. 12 (noting that } p_{rnd} \leq 1) \tag{14}$$

$$E[X_i] \;\geq\; \frac{180\max\{\log n,c\}}{\min\{2f,\frac{c}{f}\}} \;\geq\; \frac{180\max\{\log n,c\}}{\sqrt{2c}} \;>\; 90\max\{\frac{\log n}{\sqrt{c}},\sqrt{c}\} \;\geq\; 90\sqrt{\log n} \text{ (from Lemma 10)} \tag{15}$$

Note that from the preceding equations, it also proceeds that $M_u \geq \lceil \max\{\frac{45max\{\log n,c\}}{f}, 90f, 45\sqrt{\log n}\} \rceil$.

Let I_i' denote an indicator variable which is 1 if $X_i < \frac{E[X_i]}{2}$, and 0 else. Then from the Chernoff bound in Theorem 4, $Pr[I_i' = 1] = Pr[X_i < \frac{E[X_i]}{2}] \leq Pr[X_i \leq \frac{E[X_i]}{2}] \leq \exp(-\frac{E[X_i]}{8})$. Besides, the I_i''s are negatively correlated, as each node can only have f channels assigned to it, and thus, in the given set of nodes $\mathcal{E}$, having some channel (say c_i) assigned to a large number of nodes can only decrease the presence of another channel (say c_j).

Then if $X = \sum_{i=1}^{c} I_i'$, $E[X] \leq c\exp(-\frac{E[X_i]}{8}) \leq \exp(-\frac{E[X_i]}{8} + O(\log\log n)) \leq \exp(-\frac{3E[X_i]}{25})$ for large n (since $E[X_i] = \Omega(\sqrt{\log n})$ from Eqn. 15). Due to the negative correlation of I_i''s, we can still apply the Chernoff bound (Lemma 1). By setting $(1+\beta)E[X] = \frac{f}{4}$ in Theorem 2 (note that $E[X] \leq \exp(-\frac{3E[X_i]}{25}) \leq \exp(-\frac{3}{25}(180f)) < \frac{f}{4}$, yielding $\beta > 0$), we obtain by appropriate substitutions at each step, the following:

$$
\begin{aligned}
Pr[X \geq \lceil \frac{f}{4} \rceil \leq Pr[X \geq \frac{f}{4}] &\leq \left(\frac{e^{\beta}}{(1+\beta)^{(1+\beta)}} \right)^{E[X]} \\
&< \left(\frac{e}{(1+\beta)} \right)^{(1+\beta)E[X]} = \left(\frac{4eE[X]}{f} \right)^{\frac{f}{4}} \\
&\leq \left(\frac{4e\exp(-\frac{3}{25}\frac{90\max\{\log n, c\}}{f})}{f} \right)^{\frac{f}{4}} \text{ from Eqn. 13} \\
= \left(\frac{4e\exp(-\frac{270\max\{\log n, c\}}{25f})}{f} \right)^{\frac{f}{4}} &= \frac{\exp\left(\frac{-270\max\{\log n, c\}}{100}\right)}{(\frac{f}{4e})^{\frac{f}{4}}} \\
\leq \frac{\exp(-2.7\max\{\log n, c\})}{(\frac{1}{2e})^{\frac{f}{4}}} &\leq \frac{\exp(-2.7\max\{\log n, c\})}{(\frac{1}{e^2})^{\frac{f}{4}}} (\text{ since } f \geq 2) \\
&\leq \exp(-2.7\max\{\log n, c\})\exp(\frac{f}{2}) \\
&\leq \exp(-2\max\{\log n, c\}) \leq \frac{1}{n^2} (\text{ since } f \leq c)
\end{aligned}
\tag{16}
$$

Applying union bound over all $\frac{1}{a(n)} \leq n$ cells in the network, the probability that this happens in any cell is at most $\frac{1}{n}$. Thus, with probability at least $1 - \frac{1}{n}$, $X < \lceil \frac{f}{4} \rceil$, i.e., $X \leq \lfloor \frac{f}{4} \rfloor$ (since X is an integer), and hence each cell has at least $c - \lfloor \frac{f}{4} \rfloor$ channels with $X_i \geq \frac{E[X_i]}{2}$ *candidate* nodes capable of switching on them. Thus, by our definition of X, each cell has at least $c - \lfloor \frac{f}{4} \rfloor$ channels with $X_i \geq \lceil \frac{E[X_i]}{2} \rceil$ *candidate* nodes capable of switching on them (since X_i is also an integer). From Eqn. 12 and the definition of M_u, we know that $M_u = \lceil \frac{E[X_i]}{2} \rceil$. Thus, the lemma is proved. ■

Similar to the construction for connectivity from [1] that we briefly summarized in Section VI-A, we will construct a backbone for each node. *However, since our concern is not merely connectivity but also capacity, these backbones need to be constructed carefully, to ensure that no bottlenecks are formed*.

Conditioning on Lemma 11, there are at least $\frac{200\max\{\log n, c\}}{p_{rnd}}$ nodes in each cell w.h.p. Initially, from each cell, we choose $\frac{180\max\{\log n, c\}}{p_{rnd}}$ nodes uniformly at random as *backbone candidates*. The remaining nodes (which are at least $\frac{20\max\{\log n, c\}}{p_{rnd}}$ in number) are deemed *transition facilitators*.

Definition 2: (Proper Channel) A channel i is deemed *proper* in cell $\mathcal{D}$ if it occurs in at least M_u backbone candidate nodes in $\mathcal{D}$.

Lemma 17: For each cell of the network, the following is true w.h.p.: if the number of *proper* channels in the cell is c', then $c' \geq c - \lfloor \frac{f}{4} \rfloor \geq c - \lfloor \frac{c}{4} \rfloor \geq \lceil \frac{3c}{4} \rceil \geq \frac{3c}{4}$.

Proof: The proof follows from Lemma 11 and Lemma 16. ∎

Besides, we can also show the following:

Lemma 18: [2]

Consider any cell $\mathcal{D}$. Let $\mathcal{W}_i$ be the set of all nodes in the 8 adjacent cells $\mathcal{D}(k), 1 \leq k \leq 8$, that are capable of switching on channel i.

For a set of nodes $\mathcal{B}$, define $\mathcal{C}(\mathcal{B}) = \{j | j$ proper in $\mathcal{D}$ and $\exists u \in \mathcal{B}$ capable of switching on $j\}$. If $f \geq 100$, the following holds w.h.p.:

$$\forall \text{ channels i}, \forall \mathcal{B} \subseteq \mathcal{W}_i \text{ such that } |\mathcal{B}| = \lceil \frac{fna(n)}{4c} \rceil : \ |\mathcal{C}(\mathcal{B})| \geq \lceil \frac{3c}{8} \rceil$$

This is true for all cells $\mathcal{D}$.

Proof: We condition on the node-locations, and their conforming to the high-probability event of Lemma 11. Consider a cell $\mathcal{D}$. Let c' be the number of proper channels in $\mathcal{D}$.

Having conditioned on (and thus fixed) the node-locations (and thereby node-population in each cell), channel-presence in each cell is independent of other cells, as channel assignment is done independently for each node.

Then we can show that: $c' \geq c - \lfloor \frac{f}{4} \rfloor \geq c - \lfloor \frac{c}{4} \rfloor \geq \lceil \frac{3c}{4} \rceil \geq \frac{3c}{4}$, with probability at least $1 - \frac{1}{n^2}$, by following the proof argument of Lemma 16 up to Eqn. (16) (just prior to application of the union bound over all cells in the proof of that lemma).

If $c' < \frac{3c}{4}$, then we assume that our desired event does not happen for the purpose of obtaining a bound. This probability is at most $\frac{1}{n^2}$.

We now focus on the case where $c' \geq \frac{3c}{4}$.

Consider a particular channel i.

Recall that $\mathcal{W}_i$ is the set of nodes in the cells adjacent to $\mathcal{D}$ that can switch on channel i.

We first bound the probability that $|\mathcal{W}_i| \geq 2400e^2 \max\{\log n, c\}$.

Let Y_{ij} be an indicator variable that is 1 if node j in cells adjacent to $\mathcal{D}$ is capable of switching on channel i, and 0 else. Then we know that $Pr[Y_{ij} = 1] = \frac{f}{c}$, and for a given i, the Y_{ij}'s are independent. Let $Y_i = \sum_{j \in \mathcal{D}} Y_{ij}$. Then, as the node-locations conform to the high probability event of Lemma 11, $E[Y_i] \leq 8 \left(\frac{6fna(n)}{5c} \right) \leq \frac{48(250) f \max\{\log n, c\}}{5cp_{rnd}} = \frac{2400 f \max\{\log n, c\}}{cp_{rnd}}$. Setting $(1+\beta)E[Y_i] = 2400e^2 \max\{\log n, c\}$, observing from Eqn. (8) that $\beta \geq \frac{e^2 cp_{rnd}}{f} - 1 > 0$ and applying the Chernoff bound from Theorem 2:

[2]This can be viewed as a special variant of the Coupon Collector's problem [4], where there are c different types of coupons, and each box has a random subset of f different coupons. Some other somewhat different variants having multiple coupons per box have been considered in work on coding, e.g., [12].

$$Pr[Y_i \geq 2400e^2 \max\{\log n, c\}] < \left(\frac{e^{\beta}}{(1+\beta)^{(1+\beta)}}\right)^{E[Y_i]} < \left(\frac{e}{(1+\beta)}\right)^{(1+\beta)E[Y_i]}$$
$$\leq \left(\frac{fe}{e^2 c p_{rnd}}\right)^{2400e^2 \max\{\log n, c\}}$$
$$= \left(\frac{f}{ec p_{rnd}}\right)^{2400e^2 \max\{\log n, c\}} \tag{17}$$
$$\leq \left(\frac{1}{e}\right)^{2400e^2 \max\{\log n, c\}} \quad (\because \frac{f}{c p_{rnd}} \leq 1)$$
$$= \exp(-2400e^2 \max\{\log n, c\})$$
$$\leq \frac{1}{n^{2400e^2}}$$

Denote by $\mathcal{E}_{i,\mathcal{D}}$ the event that, for given i and $\mathcal{D}$: $\exists \mathcal{B} \subseteq \mathcal{W}_i'$ such that $|\mathcal{B}| = \lceil \frac{fna(n)}{4c} \rceil$ and $|\mathcal{C}(\mathcal{B})| < \lceil \frac{3c}{8} \rceil$.

Let $p_{ub}(x)$ be an upper-bound on $Pr\left[\mathcal{E}_{i,\mathcal{D}} \middle| |\mathcal{W}_i'| = x, c' \geq \frac{3c}{4}\right]$. Note that, having conditioned on (and hence fixed) the node-locations, $|\mathcal{W}_i'|$ is independent of whether $c' \geq \frac{3c}{4}$ or not.

If $p_{ub}(x)$ is an increasing function of x, then the following holds:

$$\begin{aligned}
&Pr\left[\mathcal{E}_{i,\mathcal{D}} \middle| c' \geq \frac{3c}{4}\right] \\
&= Pr\left[|\mathcal{W}_i'| \leq b | c' \geq \frac{3c}{4}\right] Pr\left[\mathcal{E}_{i,\mathcal{D}} \middle| |\mathcal{W}_i'| \leq b, c' \geq \frac{3c}{4}\right] \\
&\quad + Pr\left[|\mathcal{W}_i'| > b | c' \geq \frac{3c}{4}\right] Pr\left[\mathcal{E}_{i,\mathcal{D}} \middle| |\mathcal{W}_i'| > b, c' \geq \frac{3c}{4}\right] \\
&\leq Pr\left[|\mathcal{W}_i'| \leq b\right] Pr\left[\mathcal{E}_{i,\mathcal{D}} \middle| |\mathcal{W}_i'| \leq b, c' \geq \frac{3c}{4}\right] + Pr\left[|\mathcal{W}_i'| > b\right] \\
&= \sum_{x \leq b} Pr\left[|\mathcal{W}_i'| = x\right] Pr\left[\mathcal{E}_{i,\mathcal{D}} \middle| |\mathcal{W}_i'| = x, c' \geq \frac{3c}{4}\right] + Pr\left[|\mathcal{W}_i'| > b\right] \\
&\leq \sum_{x \leq b} Pr\left[|\mathcal{W}_i'| = x\right] p_{ub}(x) + Pr\left[|\mathcal{W}_i'| > b\right] \\
&\leq \sum_{x \leq b} Pr\left[|\mathcal{W}_i'| = x\right] p_{ub}(b) + Pr\left[|\mathcal{W}_i'| > b\right] \\
&= p_{ub}(b) \sum_{x \leq b} Pr\left[|\mathcal{W}_i'| = x\right] + Pr\left[|\mathcal{W}_i'| > b\right] \\
&= p_{ub}(b) Pr\left[|\mathcal{W}_i'| \leq b\right] + Pr\left[|\mathcal{W}_i'| > b\right] \\
&\leq p_{ub}(b) + Pr\left[|\mathcal{W}_i'| \geq b\right]
\end{aligned} \tag{18}$$

Let us now find an upper-bound $p_{ub}(x)$ that is an increasing function of x:

Note that we only need to explicitly consider $x \geq \lceil \frac{fna(n)}{4c} \rceil$, else there exist no subsets $\mathcal{B} \subseteq \mathcal{W}_i'$ satisfying $|\mathcal{B}| = \lceil \frac{fna(n)}{4c} \rceil$; thus the event $\mathcal{E}_{i,\mathcal{D}}$ cannot occur, and trivially: $p_{ub}(x) = 0$ for $0 \leq x < \lceil \frac{fna(n)}{4c} \rceil$.

If $|\mathcal{W}_i'| = x \geq \lceil \frac{fna(n)}{4c} \rceil$, then from Lemma 6, the number of subsets of $\mathcal{W}_i'$ of cardinality $m = \lceil \frac{fna(n)}{4c} \rceil$ is thus given by: $\binom{x}{m} \leq \left(\frac{xe}{m}\right)^m$.

Consider a subset $\mathcal{B} \subseteq \mathcal{W}_i'$ of specified cardinality $m = \lceil \frac{fna(n)}{4c} \rceil$. Denote by X_j the indicator variable which is 1 if channel j is not a member of $\mathcal{C}(\mathcal{B})$ and 0 else.

Recall that each node in $\mathcal{B}$ has one channel known to be i, but the remaining $f-1$ channels assigned to it are an i.i.d. chosen subset from the remaining $c-1$ available channels. Thus:

$$Pr[x \in \mathcal{W}_j (j \neq i) | x \in \mathcal{W}_i'] = \frac{f-1}{c-1} \geq \frac{f-1}{c} = \frac{f}{c}(1 - \frac{1}{f}) \geq \frac{99f}{100c} \ (\because f \geq 100) \tag{19}$$

Then from Eqn. 19, $Pr[X_j = 1] = (1 - \frac{f-1}{c-1})^{|\mathcal{B}|} \leq (1 - \frac{99f}{100c})^{\lceil \frac{fna(n)}{4c} \rceil} \leq e^{-\frac{99f}{100c} \lceil \frac{fna(n)}{4c} \rceil}$ (applying Lemma 5). Also, for a given $\mathcal{B}$, the X_j's are negatively correlated.

Let $X = \sum_{j \text{ proper in } \mathcal{D}, j \neq i} X_j$. Then $E[X] \leq c' e^{-\frac{99f}{100c} \lceil \frac{fna(n)}{4c} \rceil}$. Setting $(1+\beta)E[X] = \frac{c'}{2}$, one can see that $\beta = \frac{c'}{2E[X]} - 1 \geq \frac{c'}{2c' e^{-\frac{99f}{100c} \lceil \frac{fna(n)}{4c} \rceil}} - 1 \geq \frac{e^{\frac{99f^2 na(n)}{400c^2}}}{2} - 1 \geq \frac{e^{\frac{495}{16}}}{2} - 1 > 0$ (recall that $na(n) = \frac{250 \max\{\log n, c\}}{p_{rnd}} \geq \frac{250c \max\{\log n, c\}}{2f^2} \geq \frac{125c^2}{f^2}$, from Lemma 9). Thus we can apply the Chernoff bound from Theorem 2 to obtain that:

$$\begin{aligned}
Pr[X \geq \frac{c'}{2}] &< \left(\frac{e^{\beta}}{(1+\beta)^{(1+\beta)}} \right)^{E[X]} < \left(\frac{e}{(1+\beta)} \right)^{(1+\beta)E[X]} \\
&= \left(\frac{2eE[X]}{c'} \right)^{\frac{c'}{2}} \leq \left(\frac{2ec' \exp(-\frac{99f}{100c} \lceil \frac{fna(n)}{4c} \rceil)}{c'} \right)^{\frac{c'}{2}} \\
&= \left(2e \exp(-\frac{99f}{100c} \lceil \frac{fna(n)}{4c} \rceil) \right)^{\frac{c'}{2}} \\
&= \left(\exp(-\frac{99f}{100c} \lceil \frac{fna(n)}{4c} \rceil + (1 + \ln 2)) \right)^{\frac{c'}{2}} \\
&\left(\text{noting that } -\frac{99f}{100c} \lceil \frac{fna(n)}{4c} \rceil + (1 + \ln 2) < 0 \right) \\
&\leq \left(\exp(-\frac{99f}{100c} \lceil \frac{fna(n)}{4c} \rceil + (1 + \ln 2)) \right)^{\frac{3c}{8}} \\
&= \left(\exp(-\frac{297f}{800} \lceil \frac{fna(n)}{4c} \rceil + \frac{3c(1 + \ln 2)}{8}) \right) \\
&< \exp(-\frac{297f}{800} \lceil \frac{fna(n)}{4c} \rceil + \frac{4f}{125} \lceil \frac{fna(n)}{4c}) \rceil) \\
(\because na(n) = \frac{250 \max\{\log n, c\}}{p_{rnd}} \geq \frac{250c \max\{\log n, c\}}{2f^2}, \therefore \frac{3c(1 + \log 2)}{8} &< \frac{4f}{125} \lceil \frac{fna(n)}{4c}) \rceil) \\
&= \exp(-\frac{265f}{800} \lceil \frac{fna(n)}{4c} \rceil)
\end{aligned} \tag{20}$$

Also note that due to integrality of X, $X < \frac{c'}{2} \implies X \leq \lfloor \frac{c'}{2} \rfloor \implies |\mathcal{C}(\mathcal{B})| \geq \lceil \frac{c'}{2} \rceil \geq \lceil \frac{3c}{8} \rceil$.

Taking union bound over all possible subsets $\mathcal{B}$, we obtain that the probability it happens for any such subset $\mathcal{B}$ is at most $\left(\frac{xe}{m}\right)^m \exp(-\frac{265f}{800} \lceil \frac{fna(n)}{4c} \rceil)$ which is an increasing function of x. Thus we obtain: $p_{ub}(x) = \left(\frac{xe}{m}\right)^m \exp(-\frac{265f}{800} \lceil \frac{fna(n)}{4c} \rceil)$ for $x \geq \lceil \frac{fna(n)}{4c} \rceil$. Resultantly, $p_{ub}(x)$ is an increasing function of x.

For $b = 2400e^2 \max\{\log n, c\}$:

$p_{ub}(b) = p_{ub}(2400e^2\max\{\log n, c\}) = \left(\frac{2400e^3\max\{\log n,c\}}{\lceil\frac{fna(n)}{4c}\rceil}\right)^{\lceil\frac{fna(n)}{4c}\rceil}\exp(-\frac{265f}{800}\lceil\frac{fna(n)}{4c}\rceil) \leq \left(\frac{2400e^3\max\{\log n,c\}}{\frac{fna(n)}{4c}}\right)^{\lceil\frac{fna(n)}{4c}\rceil}\exp(-\frac{265f}{800}\lceil\frac{fna(n)}{4c}\rceil) \leq \left(\frac{9600e^3 cp_{rnd}}{250f}\right)^{\lceil\frac{fna(n)}{4c}\rceil}\exp(-\frac{265f}{800}\lceil\frac{fna(n)}{4c}\rceil) \leq \exp((3+\log\frac{960}{25}+\log\frac{cp_{rnd}}{f})\lceil\frac{fna(n)}{4c}\rceil)\exp(-\frac{265f}{800}\lceil\frac{fna(n)}{4c}\rceil) < \exp((3+\log 40+\log 2f)\lceil\frac{fna(n)}{4c}\rceil)\exp(-\frac{265f}{800}\lceil\frac{fna(n)}{4c}\rceil)$ (using Lemma 9).

Since $f \geq 100$, the following always holds: $f \geq 8(3+\log 40+\log 2f)$. Thus $p_{ub}(b) \leq \exp(\frac{f}{8}\lceil\frac{fna(n)}{4c}\rceil)\exp(-\frac{265f}{800}\lceil\frac{fna(n)}{4c}\rceil) = \exp(-\frac{165f}{800}\lceil\frac{fna(n)}{4c}\rceil) < \exp(-\frac{f}{5}\lceil\frac{fna(n)}{4c}\rceil) < \exp(-\frac{f^2na(n)}{20c}) \leq \exp(-\frac{125\log n}{20}) < \frac{1}{n^6}$ (from Lemma 9 and our choice of $a(n)$).

Thus from Eqn. (18), $Pr[\mathcal{E}_{i,\mathcal{D}}|c' \geq \frac{3c}{4}] \leq p_{ub}(b) + Pr[|\mathcal{W}_i| \geq b] \leq \frac{1}{n^6} + \frac{1}{n^{2400e^2}} < \frac{1}{n^5}$.

Since there are $c = O(\log n)$ channels i to consider, we take a union bound over them to obtain that: $Pr[\mathcal{E}_{i,\mathcal{D}}$ for any i in $\mathcal{D}|c' \geq \frac{3c}{4}] \leq cPr[\mathcal{E}_{i,\mathcal{D}}$ for a given i in $\mathcal{D}|c' \geq \frac{3c}{4}])$.

Thus: $Pr[\mathcal{E}_{i,\mathcal{D}}$ for any i in $\mathcal{D}$ $] \leq Pr[c' < \frac{3c}{4}] + Pr[c' \geq \frac{3c}{4}](cPr[\mathcal{E}_{i,\mathcal{D}}$ for a given i in $\mathcal{D}|c' \geq \frac{3c}{4}]) \leq Pr[c' < \frac{3c}{4}] + cPr[\mathcal{E}_{i,\mathcal{D}}$ for a given i in $\mathcal{D}|c' \geq \frac{3c}{4}] \leq \frac{1}{n^2} + \frac{c}{n^5}$

We then take another union bound over all $\frac{1}{a(n)} = \frac{p_{rnd}n}{250\max\{\log n,c\}} < \frac{n}{c}$ cells $\mathcal{D}$ to obtain that the probability this occurs in any cell is at most $\frac{1}{cn} + \frac{1}{n^4}$.

Finally, recall that we conditioned our proof on the node-locations conforming to the high-probability event of Lemma 11. The probability that this event does not occur is at most $\frac{50\log n}{n}$ (as proved in Lemma 11), and we can obtain a bound by assuming that whenever that event fails to hold, the event in the statement of this lemma fails to hold.

This completes the proof that $\mathcal{C}(\mathcal{B}) \geq c' - \lfloor\frac{c'}{2}\rfloor \geq \lceil\frac{c'}{2}\rceil \geq \lceil\frac{3c}{8}\rceil$ for all specified subsets $\mathcal{B}$ of interest, for all channels i, and in all cells $\mathcal{D}$ with probability at least $1 - \frac{1}{cn} - \frac{1}{n^4} - \frac{50\log n}{n} > 1 - \frac{2}{n} - \frac{50\log n}{n}$.

■

A. Routing and channel assignment

Partial Backbones: As mentioned earlier, the routing strategy is based on a per-node backbone structure similar to that used to prove the sufficient condition for connectivity. However, *instead of constructing a full backbone for each node, only a partial backbone $\mathcal{B}_p(x)$ is constructed for each node x.* $\mathcal{B}_p(x)$ only covers those cells which are traversed by flows for which x is either source or destination. A flow first proceeds along the route on the source backbone and will then attempt to switch onto the destination backbone.

We shall explain the backbone construction procedure in detail later. First we show how a flow can be routed along these backbones from its source to its destination.

Lemma 19: Suppose a flow has source x and destination y. Thus it is initially on $\mathcal{B}_p(x)$ and finally needs to be on $\mathcal{B}_p(y)$. Then after having traversed $\frac{c^2}{f^2}$ distinct cells (hops) (recall that $2 \leq f \leq c$ and $c = O(\log n)$) , it will have found an opportunity to make the transition w.h.p. If the routes of each of the n flows get to traverse at least $\frac{c^2}{f^2}$ distinct cells (note that each individual route needs to traverse at least so many distinct cells; two different flows may share cells on their respective routes), then all n flows are able to transition w.h.p.

Proof: Consider a flow traversing a sequence of cells $D_1, D_2, \ldots$. Then if the representative of $\mathcal{B}_p(x)$ (let us call it q_x) in D_i can communicate (directly or indirectly) with the representative of $\mathcal{B}_p(y)$ (let us call it q_y) in D_i, it is possible to switch from $\mathcal{B}_p(x)$ to $\mathcal{B}_p(y)$. If q_x and q_y share a channel this is trivial. If q_x and q_y do not share a channel, we consider the probability that the two can communicate via a third node from amongst the *transition facilitators* in D_i, i.e. there exists a transition facilitator z such that z shares at least one channel with q_x and one channel with q_y. In Section VI-A, we summarized a proof from [1] showing that q_x and q_y can communicate through a given z with probability $p_z > \frac{f^2}{c^2} = \Omega(\frac{1}{\log^2 n})$. Given our choice of cell area $a(n)$, and conditioned on the fact that each cell has $\frac{200\max\{\log n, c\}}{p_{rnd}}$ nodes (Lemma 11), of which $\frac{180\max\{\log n, c\}}{p_{rnd}}$ are deemed *backbone candidates* and the rest are *transition facilitators*, there are at least $20\frac{\max\{\log n, c\}}{p_{rnd}} \geq \frac{20\log n}{p_{rnd}}$ possibilities for z within that cell. All the possible z nodes have i.i.d. channel assignments. Thus, the probability that q_x and q_y cannot communicate through any z in the cell is at most $(1-p_z)^{\frac{20\log n}{p_{rnd}}}$, and the probability they communicate through some z is $p_{xy} \leq 1-(1-p_z)^{\frac{20\log n}{p_{rnd}}}$.

Hence, the probability that this happens in none of the $\frac{c^2}{f^2}$ distinct cells is at most $(1-p_{xy})^{\frac{c^2}{f^2}} < (1-p_z)^{\frac{20c^2\log n}{f^2 p_{rnd}}} < (1-\frac{f^2}{c^2})^{\frac{20c^2\log n}{f^2 p_{rnd}}} \leq e^{-\frac{20\log n}{p_{rnd}}} \leq \frac{1}{n^{20}}$ (from Lemma 5). Applying union bound over all n flows, the probability that all flows are able to transition is at least $1-\frac{1}{n^{19}}$. ■

Therefore, we require each route to have at least $\frac{c^2}{f^2}$ distinct hops [3]. Resultantly, we cannot stipulate that *all* flows be routed along the (almost) straight-line path $SD'D$ (Fig. 1). If $SD'D$ is short, a detour may be required to ensure the minimum route-length, akin to detour-routing in the constructions of [1]. Such flows are said to be *detour-routed*.

Flow Transition Strategy: As per our strategy, a non-detour-routed flow is initially in a *progress-on-source-backbone* mode, and keeps to the source backbone till there are only $\frac{c^2}{f^2}$ distinct intermediate cells left to the destination. At this point, it enters a *ready-for-transition* mode, and actively seeks opportunities to make a transition to the destination backbone along the remaining hops. Once it has made the transition into the destination backbone, it proceeds towards the destination on that backbone along the remaining part of the route, and is thus guaranteed to reach the destination.

Thus, we stipulate that the (almost) straight-line $SD'D$ path be followed if the straight-line route comprises $h \geq \frac{c^2}{f^2}$ distinct intermediate cells (hops). If S and D' (hence also D) lie close to each other, the hop-length of the straight line cell-to-cell path can be much smaller. In this case, a *detour* path $SPD'D$ is chosen (Fig. 2), using a circle of radius $\frac{c^2}{f^2}r(n)$ in a manner similar to that in the constructions described in [1], [9] (consider a circle of this radius centered around S, choose a point P on the circle, and follow the route $SPD'D$).

A detour-routed flow is always in *ready-for-transition* mode.

The need to perform *detour* routing for some source-destination pairs does not have any substantial effect on the average hop-length of routes or the relaying load on a cell, as we show further.

Lemma 20: The length of any route increases by at most $O(\log^2 n)$ hops w.h.p.

Proof: The proof proceeds directly from the *detour* routing strategy. Recall that the area of a cell is $\frac{250\max\{\log n, c\}}{p_{rnd}n}$, i.e., the side of each cell is $\Theta(\sqrt{\frac{\log n}{p_{rnd}n}})$ (more precisely it is $\frac{r(n)}{\sqrt{8}}$). The distance SP in Fig. 2 is

[3] Note that this does not constitute a tight bound on the minimum number of hops required for a transition.

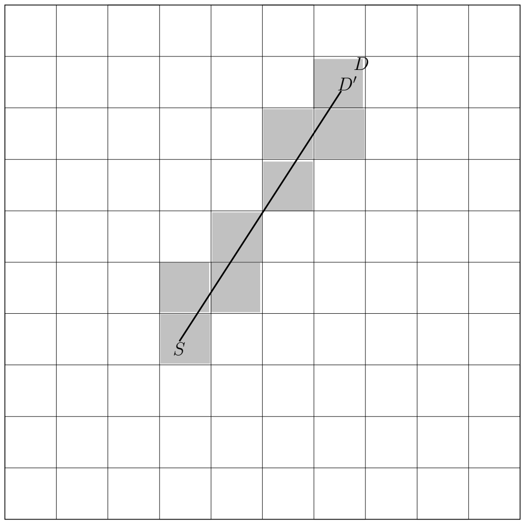

Fig. 1. Routing along a straight line

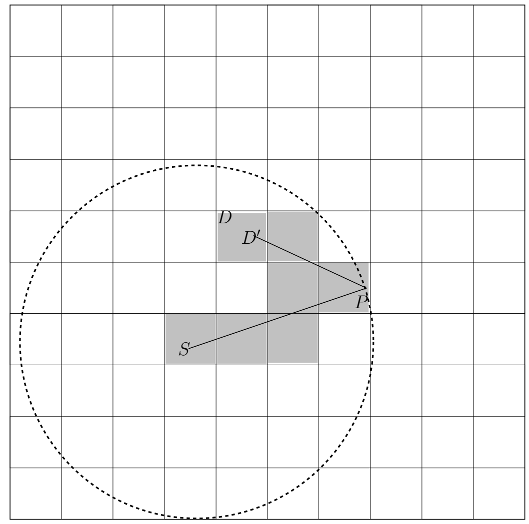

Fig. 2. Illustration of detour routing

at most $\frac{c^2}{f^2}r(n)$ (radius of the circle in the figure), yielding at most $O(\frac{c^2}{f^2})$ hops, while PD is again *at most* $\Theta(\frac{c^2}{f^2})$ hops (diameter of circle). This increases route length by at most $O(\frac{c^2}{f^2}) \implies O(\log^2 n)$ hops (recall that $c = O(\log n)$). ■

Lemma 21: If the number of flows in any cell is x in case of pure straight-line routing, it is at most $x + O(\frac{nc^4r^2(n)}{f^4}) \implies x + O(\log^6 n)$ w.h.p. in case of detour routing.

Proof: Recall that $c = O(\log n)$. Since the detour occurs only up to a circle of radius $\frac{c^2}{f^2}r(n)$, the extra flows that may pass through a cell (compared to straight-line routing) are only those whose sources lie within a distance $\frac{c^2}{f^2}r(n)$ from some point in this cell. Thus all such possible sources fall within a circle of radius $(1+\frac{c^2}{f^2})r(n)$, and hence area $a_c(n) = \Theta(\frac{c^4r^2(n)}{f^4})$. Applying Lemma 3 to the set of n node locations (with a suitable choice of $\alpha(n) \geq 1$), with high probability, any circle of this radius will have $O(na_c(n))$ nodes, and hence $O(na_c(n))$ sources. Hence the number of extra flows that traverse the cell due to detour routing is $O(na_c(n))$, and each detour-routed flow can traverse a cell at most twice. Thus, the total number of flows (even counting repeat traversals separately) $x + O(\frac{nc^4r^2(n)}{f^4})$. Since $nr^2(n) = O(\frac{\log n}{p_{rnd}})$, and $p_{rnd} \geq \frac{f}{c}$, the total number of flows is $O(\frac{c^5 \log n}{f^5}) \implies x + O(\log^6 n)$ w.h.p. ■

Lemma 22: The number of flows traversing any cell is $O(n\sqrt{a(n)})$ w.h.p. even with detour routing.

Proof: From Lemma 15, we know that the number of flows passing through a cell with $SD'D$ routing (without detours) is $O(n\sqrt{a(n)})$. Thus, from Lemma 21, the number of flows through the cell , even after some flows are detour-routed, is at most $O(n\sqrt{a(n)}) + O(\log^6 n) \implies O(n\sqrt{a(n)})$ (since $a(n) = \Theta(\frac{\log n}{p_{rnd}})$). ■

Lemma 23: The number of flows traversing any cell in *ready-for-transition* mode is $O(\log^6 n)$ w.h.p.

Proof: First let us account for the SD' stretch of each flow, without considering the possible additional last hop. We account for it explicitly later in this proof.

By our construction, a non-detour routed flow enters the *ready-for-transition* mode only when it is $\frac{c^2}{f^2}$ hops away from its destination. All such flows must have their pseudo-destinations within a circle of radius $\Theta(\frac{c^2}{f^2}r(n))$ centered in the cell. The number of pseudo-destinations that lie within a circle of radius $\Theta(\frac{c^2}{f^2}r(n))$ from the cell is $\Theta(\frac{nc^4r^2(n)}{f^4}) \implies O(\frac{c^5}{f^5}\log n)$ w.h.p., (by observing that $p_{rnd} \geq \frac{f}{c}$, and using suitable choice of $\alpha(n) = O(\frac{c^5}{f^5})$ in Lemma 3). Also $c = O(\log n)$. Hence there are $O(\log^6 n)$ non-detour-routed flows in *ready-for-transition* mode traversing the cell w.h.p.

A detour-routed flow is always in *ready-for-transition* mode. By Lemma 21, there are $O(\log^6 n)$ such flows traversing any cell. Each such flow can only traverse a cell twice along the SD' stretch. This yields $O(\log^6 n)$ detour-routed flows (including repeat traversals).

Also, the cell may be re-traversed by some flows on their additional last hop. From Lemma 14, the pseudo-destinations of such flows must lie in the same cell or one of the 8 adjacent cells. Applying Lemma 2 to the set of n pseudo-destinations, it proceeds that the total number of pseudo-destinations lying in these 9 cells is $\Theta(na(n))$ w.h.p. Thus, the number of flows entering the cell on their additional last hop is $O(na(n)) \implies O(\log^2 n)$.

Hence the number of flows transitioning in any cell is $O(\log^6 n)$ w.h.p.

■

Backbone Construction: The backbone construction procedure is required to take load-balancing into account. Thus we can describe the procedure for constructing the backbone $\mathcal{B}_p(x)$ of x as follows:

Given a cell $\mathcal{D}$, the 8 cells adjacent to cell $\mathcal{D}$ are denoted as $\mathcal{D}(j), 1 \leq j \leq 8$ (Fig. 3). $\mathcal{B}_p(x)$ is constituted as follows. Let $\mathcal{S} \cup \mathcal{D}_b$ be the subset of cells that must be covered by $\mathcal{B}_p(x)$ where $\mathcal{S}$ comprises cells traversed by the flow for which x is the source, and $\mathcal{D}_b$ comprises the cells traversed by flows for which it may be the destination. x is by default a member of $\mathcal{B}_p(x)$.

We consider backbone construction for the route from each source to its pseudo-destination below. Some routes require an additional last hop to reach the actual destination node. However, from Lemma 14, the only such last hop routes that may enter a cell correspond to pseudo-destinations in the 8 adjacent cells. Then applying Lemma 2 to the set of pseudo-destinations, they are only $O(na(n))$ such pseudo-destinations, and thus only $O(na(n))$ such last-hop flows entering the cell. Hence we can account for them separately.

a) Expanding backbones to $\mathcal{S}$: We first cover cells in $\mathcal{S}$. Recall that we are only constructing the SD' part and not considering the possible additional last hop at this stage.

This has two sub-stages. In the first stage, we construct backbones for source nodes whose flow does not require a detour. In the second sub-stage we construct backbones for source nodes whose flow requires a detour.

Straight-line backbones:

This step proceeds in a hop-by-hop manner for all non-detour-routed flows in parallel (each of which has a unique source x).

Any cell of $\mathcal{S}$ in which there is already a node assigned to $\mathcal{B}_p(x)$ is called a filled cell. Thus initially x's cell is

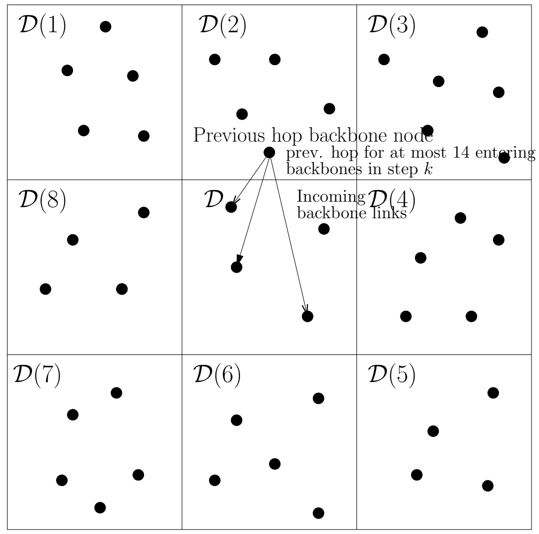

Fig. 3. Cell $\mathcal{D}$ and neighboring cells during backbone construction

filled. We then consider the cell in $\mathcal{S}$ that is traversed next by the flow. We consider all nodes in that cell sharing one or more common channel with x. This provides a number of alternative channels on which the flow can enter that cell.

Let h_{max} be the maximum hop-length of any non-detour-routed SD' route. Then, $h_{max} = O(\frac{1}{\sqrt{a(n)}})$ and the procedure has h_{max} steps. In step k, for each source node x whose flow has k or more hops, $\mathcal{B}_p(x)$ expands into the cell entered by x's flow on the k-th hop. Each cell $\mathcal{D}$ performs the following procedure:

The backbones are extended by constructing bipartite graphs that aid load-balance.

Lemma 24: If $f \geq 100$, then it is possible to devise a backbone construction procedure, such that, after step h_{max} of the backbone construction procedure for $\mathcal{S}$ (for non-detour-routed flows), each cell has $O(\frac{n\sqrt{a(n)}}{c})$ incoming backbone links on a single channel, and each node appears on $O(\frac{n\sqrt{a(n)}}{c})$ (source) backbones, w.h.p.

Proof: This proof assumes the high probability events in Lemma 11, Lemma 12, Lemma 17, and Lemma 18 occur.

We present an inductive argument. Recall that we are expanding backbones to cover cells in $\mathcal{S}$. *At each step of the (inductive) construction, we first have a channel-allocation phase, followed by a node-allocation phase*. We prove that after step k of the backbone construction procedure, the following two invariants hold for *all* cells of the network:

- *Invariant 1:* Each node is assigned at most 14 new incoming backbone links during step k. Thus after step k, it appears in a total of $O(14k) \implies O(k)$ backbones.
- *Invariant 2:* No more than $\lfloor \frac{5na(n)}{c} \rfloor$ new backbone links enter the cell on a single channel during step k. Thus, in total $O(\frac{kna(n)}{c})$ incoming backbones (entering the cell) are assigned (incoming links) on a single channel after step k.

If the above two Invariants hold, then it is easy to see that after h_{max} steps, cell $\mathcal{D}$ will have no more than

$\frac{5h_{max}na(n)}{c} = O(\frac{n\sqrt{a(n)}}{c})$ backbone links assigned to any single channel, and no node occurs on more than $14h_{max} \implies O(\frac{1}{\sqrt{a(n)}}) \implies O(\frac{n\sqrt{a(n)}}{c})$ backbones (from Eqn. (10)).

We prove that the Invariants hold, by induction, as follows:

If Invariant 1 holds at the end of step $k-1$, then Invariant 2 continues to hold after the channel-allocation phase of step k. If Invariant 2 holds after the channel-allocation phase of step k, then Invariant 1 will continue to hold after the node-allocation phase of step k, and thus both Invariants 1 and 2 will hold at the end of step k.

Base Case:

Before the procedure begins, at step 0, each node is assigned to its own backbone, for which it is effectively the origin (and this can be viewed as a single backbone link incoming to this node from an imaginary super-source). Thus after Step 0, Invariant 1 holds trivially, and Invariant 2 is irrelevant, and thus trivially true.

Inductive Step:

Suppose Invariants 1 and 2 held at the end of step $k-1$. Consider a particular cell $\mathcal{D}$ during step k.

Let the number of *proper* channels in $\mathcal{D}$ be c'. From Lemma 17, we know that $c' \geq c - \lfloor \frac{f}{4} \rfloor \geq \frac{3c}{4}$ for each cell. Each flow that enters cell $\mathcal{D}$ in step k has a previous hop-node in one of the 8 adjacent cells. Also note that, from Lemma 17, each previous hop node has at least $\lceil \frac{3f}{4} \rceil$ of cell $\mathcal{D}$'s *proper* channels available to it as choices (since it has f channels of which at most $\lfloor \frac{f}{4} \rfloor$ may be non-proper in cell $\mathcal{D}$).

Channel-Allocation: Construct a bipartite graph with two sets of vertices (Fig. 4); one set (call it $\mathcal{L}$) has a vertex corresponding to each of the (source) backbones that enter the cell $\mathcal{D}$ in step k. From Lemma 12, it proceeds that $|\mathcal{L}| \leq \lfloor \frac{5na(n)}{4} \rfloor$. The other set (call it $\mathcal{P}$) has $\lfloor \frac{5na(n)}{c} \rfloor \leq \frac{5na(n)}{c}$ vertices for each proper channel i in cell $\mathcal{D}$, i.e., $|\mathcal{P}| = c' \lfloor \frac{5na(n)}{c} \rfloor$.

A backbone vertex is connected to all the vertices for the channels proper in $\mathcal{D}$ on which its previous hop node can switch (and which are therefore valid channel choices for entering the cell $\mathcal{D}$). We show that there exists a matching that pairs each backbone vertex to a unique channel vertex, through an argument based on Hall's marriage theorem (Theorem 5). Thus, we seek to show that for all $\mathcal{V} \subseteq \mathcal{L}$, $|\mathcal{N}(\mathcal{V})| \geq |\mathcal{V}|$, where $\mathcal{N}(\mathcal{V}) \subseteq \mathcal{P}$ is the union of the neighbor-sets of all vertices in $\mathcal{V}$.

We first note the following:

$$\begin{aligned} \lceil \frac{3f}{4} \rceil \lfloor \frac{5na(n)}{c} \rfloor \geq \frac{3f}{4} \left(\frac{5na(n)}{c} - 1 \right) = \frac{15fna(n)}{4c} - \frac{3f}{4} \\ \geq \frac{15fna(n)}{4c} - \frac{3fna(n)}{1000c} \geq \frac{29fna(n)}{8c} (\because na(n) \geq 250c) \end{aligned} \tag{21}$$

Consider the following two cases:

Case 1: $|\mathcal{V}| < \frac{29fna(n)}{8c}$: Consider any set $\mathcal{V}$ of backbone vertices such that $|\mathcal{V}| < \frac{29fna(n)}{8c}$. Then, since there are at most $\lfloor \frac{f}{4} \rfloor$ non-proper channels in a cell, every previous hop node has at least $\lceil \frac{3f}{4} \rceil \geq \frac{3f}{4}$ *proper* channel choices. For each proper channel there are $\lfloor \frac{5na(n)}{c} \rfloor \geq \frac{5na(n)}{c} - 1$ associated channel vertices. Thus we obtain that

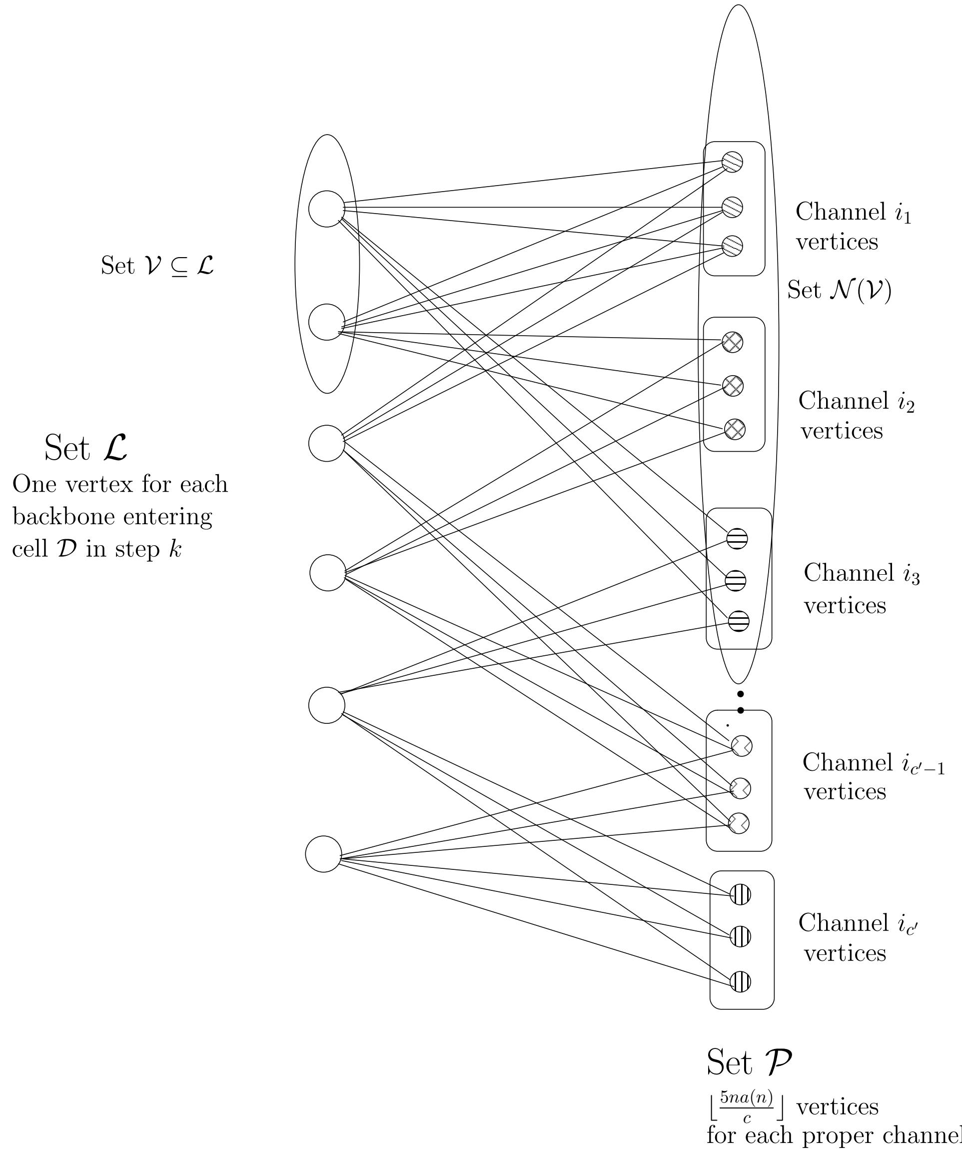

Fig. 4. Bipartite Graph for Cell $\mathcal{D}$ in step k

$|\mathcal{N}(\mathcal{V})| \geq \frac{3f}{4}\left(\frac{5na(n)}{c} - 1\right) \geq \frac{29fna(n)}{8c}$ (from Eqn. 21). Thus $|\mathcal{N}(\mathcal{V})| \geq |\mathcal{V}|$.

Case 2: $|\mathcal{V}| \geq \frac{29fna(n)}{8c}$*:* Now consider sets $\mathcal{V}$ of size at least $\frac{29fna(n)}{8c}$. Note that since Invariant 1 held till end of step $k-1$, no more than 14 backbone links were assigned to any single node in $\bigcup_{k=1}^{8} \mathcal{D}(k)$ in step $k-1$.

Intuitively, in order to show that $|\mathcal{N}(\mathcal{V})| \geq |\mathcal{V}|$ for all such $\mathcal{V}$, we first state and prove the observation that if a channel overload condition occurs, resulting in $|\mathcal{N}(\mathcal{V})| < |\mathcal{V}|$ for some $\mathcal{V}$, then the overload must also manifest itself in some *channel-aligned* subset (i.e. a subset where all flows have some *common* proper channel i available to them). Thus, to show that no overload condition occurs, it suffices to show that no overload condition occurs in any of these *critical* channel-aligned subsets, which can be shown using Lemma 18. The argument is formalized as follows:

Let $\mathcal{V}_i$ be the set comprising all sets $\mathcal{U}_i \subseteq \mathcal{L}$, such that all backbone vertices in $\mathcal{U}_i$ have channel i associated with them (i.e., all backbone vertices in $\mathcal{U}_i$ have i available to them as a valid proper channel choice for entering $\mathcal{D}$).

Claim (a): $\forall \mathcal{U} \in \bigcup_{i \text{ proper in } \mathcal{D}} \mathcal{V}_i :$

$$|\mathcal{U}| \geq \lceil \frac{29fna(n)}{8c} \rceil \implies |\mathcal{N}(\mathcal{U})| \geq |\mathcal{L}|$$

Proof of Claim (a): We know that $\mathcal{U} \in \mathcal{V}_i$ for some i that is proper in $\mathcal{D}$. Also, since no node can be the previous hop in step k of more flows than those assigned to it in step $k-1$, and Invariant 1 held after step $k-1$, it proceeds that no previous hop node is common to more than 14 entering backbone links. Let $\mathcal{A}$ be the set of distinct previous hop nodes associated with $\mathcal{U}$. Then $|\mathcal{A}| \geq \frac{1}{14}|\mathcal{U}| \geq \frac{1}{14}(\frac{29fna(n)}{8c}) \geq \frac{fna(n)}{4c} + \frac{fna(n)}{112c} > \frac{fna(n)}{4c} + 1 \geq \lceil \frac{fna(n)}{4c} \rceil$ (note that $\frac{fna(n)}{c} \geq 250f \geq 500 > 112$). Observe that $\mathcal{A}$ thus contains at least one subset $\mathcal{B}$ satisfying $|\mathcal{B}| = \lceil \frac{fna(n)}{4c} \rceil$. Recognizing that all members of $\mathcal{A}$, and hence all members of $\mathcal{B}$, are capable of switching on channel i, we can invoke Lemma 18 on $\mathcal{B}$, to obtain that when $f \geq 100$: $|\mathcal{C}(\mathcal{B})| \geq \lceil \frac{3c}{8} \rceil$. This yields: $\mathcal{N}(\mathcal{U}) \geq |\mathcal{C}(\mathcal{B})| \lfloor \frac{5na(n)}{c} \rfloor \geq |\mathcal{C}(\mathcal{B})| \left(\frac{5na(n)}{c} - 1 \right) \geq \lceil \frac{3c}{8} \rceil \left(\frac{5na(n)}{c} - 1 \right) \geq \frac{15na(n)}{8} - \lceil \frac{3c}{8} \rceil \geq \frac{15na(n)}{8} - \frac{3}{8} \left(\frac{na(n)}{250} \right) - 1 \geq \frac{5na(n)}{4} \geq |\mathcal{L}|$.

Claim (b): Consider a set $\mathcal{V} \subseteq \mathcal{L}$. Then:

$$|\mathcal{N}(\mathcal{V})| < |\mathcal{V}| \implies \exists i \text{ proper in } \mathcal{D}, \mathcal{S}_i \subseteq \mathcal{V} \text{ s.t. } : \\ \mathcal{S}_i \in \mathcal{V}_i \text{ and } |\mathcal{S}_i| \geq \lceil \frac{29fna(n)}{8c} \rceil \tag{22}$$

Proof of Claim (b): Suppose $|\mathcal{N}(\mathcal{V})| < |\mathcal{V}|$. Let us denote by $\mathcal{S}_i \subseteq \mathcal{V}$ the set of all backbone vertices in $\mathcal{V}$ that are associated with channel i (i.e., have channel i available as a valid proper channel choice for entering cell $\mathcal{D}$). Consider the bipartite sub-graph $G_{\mathcal{V}}$ induced by $\mathcal{V} \cup \mathcal{N}(\mathcal{V})$, and assign all edges unit capacity. Construct the graph $G_{\mathcal{V}} \cup \{s,t\}$ where s is a source node having a unit capacity edge to all vertices $v \in \mathcal{V}$, and t is a sink node, connected to each vertex $u \in \mathcal{N}(\mathcal{V})$ via a unit capacity edge. We try to obtain a (s,t) flow g such that all edges (s,v) are saturated. Each vertex $v \in \mathcal{V}$ sub-divides the unit of flow received from s equally amongst all edges (v,u) outgoing from it. Since each vertex has edges to vertices of at least $\frac{3f}{4}$ channels, this yields at least $\frac{3f}{4} \left(\frac{5na(n)}{c} - 1 \right) \geq \frac{29fna(n)}{8c}$ edges (see Eqn. 21). Thus each $v \in \mathcal{V}$ contributes at most $\frac{8c}{29fna(n)}$ units of flow to a vertex $u \in \mathcal{N}(\mathcal{V})$, i.e., $g(v,u) \leq \frac{8c}{29fna(n)}$. Hence no vertex $u \in \mathcal{N}(\mathcal{V})$ gets more than $h(u) = \sum_{v \in \mathcal{S}_i} g(v,u) = \frac{8c|\mathcal{S}_i|}{29fna(n)}$ units of flow, where i is the channel corresponding to vertex u. Resultantly, if $|\mathcal{S}_i| \leq \lfloor \frac{29fna(n)}{8c} \rfloor$ for all channels i that are proper in cell $\mathcal{D}$, this implies that $h(u) \leq 1$, and setting $g(u,t) = h(u)$ yields the desired (s,t) flow. Hence g is a valid flow that allows a unit of flow to pass through each vertex $v \in \mathcal{V}$. From the Integrality Theorem (Theorem 6), we can obtain an integer-capacity flow that yields a matching of size $|\mathcal{V}|$. Thus, from Hall's marriage theorem (Theorem 5), $|\mathcal{N}(\mathcal{V})| \geq |\mathcal{V}|$ (else a matching of size $|\mathcal{V}|$ could not have existed). This yields a contradiction. Thus there must exist a proper channel i, and $\mathcal{S}_i \subseteq \mathcal{V}$ such that $\mathcal{S}_i \in \mathcal{V}_i$ and $|\mathcal{S}_i| > \lfloor \frac{29fna(n)}{8c} \rfloor$. Since set-cardinality must necessarily be an integer, it proceeds that $|\mathcal{S}_i| \geq \lceil \frac{29fna(n)}{8c} \rceil$, and Eqn. (22) holds.

Claim (c): $\forall \mathcal{V} \subseteq \mathcal{L}$ such that $|\mathcal{V}| \geq \frac{29fna(n)}{8c} : |\mathcal{N}(\mathcal{V})| \geq |\mathcal{V}|$

Proof of Claim (c): Suppose $|\mathcal{N}(\mathcal{V})| < |\mathcal{V}|$. Then, from Claim (b), there exists a set $\mathcal{S}_i \subseteq \mathcal{V}$ such that $\mathcal{S}_i \in \mathcal{V}_i$, and $|\mathcal{S}_i| \geq \lceil \frac{29fna(n)}{8c} \rceil$. Thus $\mathcal{S}_i$ qualifies as a set to which Claim (a) applies. Invoking Claim (a) on this set $\mathcal{S}_i$, it follows that $|\mathcal{N}(\mathcal{V})| \geq |\mathcal{N}(\mathcal{S}_i)| \geq |\mathcal{L}| \geq |\mathcal{V}|$. This yields a contradiction. Thus, $|\mathcal{N}(\mathcal{V})| \geq |\mathcal{V}|$.

Hence, by application of Hall's marriage theorem (Theorem 5), each backbone vertex can be matched with a unique channel vertex, and the corresponding backbone will be assigned to the channel with which this vertex is associated. Thus all backbones get assigned a channel, and (since there are $\lfloor \frac{5na(n)}{c} \rfloor$ channel vertices for each proper channel) no more than $\lfloor \frac{5na(n)}{c} \rfloor$ incoming backbone links are assigned to any single channel.

While Hall's marriage theorem proves that such a matching exists, the matching itself can be computed using the Ford-Fulkerson method [2] on a flow network obtained from the bipartite graph by adding a source with an edge to each vertex in $\mathcal{L}$, a sink to which each vertex in $\mathcal{P}$ has an edge, and assigning unit capacity to all edges.

Thus Invariant 2 continues to hold after the channel-allocation phase of step k. [4]

Node-Allocation: Having determined the channel each backbone should use to enter cell $\mathcal{D}$, we need to assign a node in cell $\mathcal{D}$ to each backbone. For this, we again construct a bipartite graph. In this graph, the first set of vertices (call it $\mathcal{F}$) comprise a vertex for each backbone entering cell $\mathcal{D}$ in step k. The second set (call it $\mathcal{R}$) comprises 14 vertices for each *backbone candidate* node in cell $\mathcal{D}$. A vertex x in $\mathcal{F}$ has an edge with a vertex y in $\mathcal{R}$ iff the actual *backbone candidate* node associated with y is capable of switching on the channel assigned to the backbone associated with vertex x in the preceding channel-allocation phase.

Each vertex $x \in \mathcal{F}$ has degree at least $14M_u$, since it is assigned to a *proper* channel, which by definition has at least M_u representatives in cell $\mathcal{D}$, each of which has 14 associated vertices in $\mathcal{R}$. Also recall that $M_u = \lceil \frac{9fna(n)}{25c} \rceil$. Once again we seek to show that for all $\mathcal{V} \subseteq \mathcal{F}, |\mathcal{N}(\mathcal{V})| \geq |\mathcal{V}|$.

Consider any set $\mathcal{V} \in \mathcal{F}$.

Since no channel is assigned more than $\lfloor \frac{5na(n)}{c} \rfloor$ entering backbone links in this step, the vertices in $\mathcal{V}$ are cumulatively associated with at least $m \geq \frac{|\mathcal{V}|}{\lfloor \frac{5na(n)}{c} \rfloor}$ distinct proper channels. Since each of these channels have at least M_u *backbone candidate* nodes capable of switching on them, and any one node can only switch on up to f proper channels, this implies that the number of nodes in cell $\mathcal{D}$ cumulatively associated with these $m \geq \frac{|\mathcal{V}|}{\lfloor \frac{5na(n)}{c} \rfloor}$

[4] It is interesting to consider whether load-balance would continue to hold even if we follow simpler procedures. We have shown in [1], [9] that for random (c, f) assignment, a per-flow throughput of $\Theta(W\sqrt{\frac{f}{cn\log n}})$ is achievable with a much simpler construction. That construction is of interest despite not achieving optimal capacity since it provides a trade-off between throughput and routing/scheduling complexity. In fact when f is a small constant, the asymptotic capacity for both constructions is within a small constant factor of each other. However, it is also useful to consider whether simpler procedures can allow one to achieve the optimal capacity. As an illustration, consider a procedure where a backbone link is assigned to the least-loaded of all channels available to it. If this procedure can be proved to yield optimal load-balance, it would have useful practical implications toward potentially indicating that even simple protocols can suffice for good performance. This problem is a special variant of the problem of throwing balls into bins with the power of d choices. The problem of throwing a balls into b bins with d choices was studied in [13]. In [14], a balls-and-bins technique is used to obtain fractional matchings in graphs. However these results yield probability bounds polynomial in number of bins. In our case, the bins (channels) are $O(\log n)$ (where n is number of nodes), and we need much stronger bounds to ensure that global overload probability goes to 0, and thus a simple adaptation of existing balls-into-bins proofs does not suffice. Our case also has additional constraints, e.g., the number of choices available to each ball is $\Theta(f)$, and the number of balls (traversing source backbones) decreases with increase in f.

Also of interest is the possibility of having optimal-capacity achieving procedures where backbones are constructed sequentially, or even better, completely asynchronously (recall that the simpler construction possesses these properties, but yields sub-optimal capacity). If such a procedure can be shown to achieve good load balance, it has useful protocol implications in that when a new flow is admitted, routes for existing flows do not need to be re-organized to ensure load-balance.

proper channels is at least $\frac{|\mathcal{V}|M_u}{f\lfloor\frac{5na(n)}{c}\rfloor} \geq \frac{|\mathcal{V}|\lceil\frac{9fna(n)}{25c}\rceil}{\frac{5fna(n)}{c}} \geq \frac{9|\mathcal{V}|}{125}$, and as each node has 14 vertices, it follows that $|\mathcal{N}(\mathcal{V})| \geq 14\left(\frac{9|\mathcal{V}|}{125}\right) \geq \frac{126|\mathcal{V}|}{125} > |\mathcal{V}|$.

Then invoking Hall's Marriage Theorem again, each vertex $x \in \mathcal{F}$ can be matched with a unique vertex $y \in \mathcal{R}$, and the actual network node associated with y is deemed the backbone representative for the backbone corresponding to vertex x in cell $\mathcal{D}$ (the matching can again be computed via the Ford-Fulkerson method). Since there are at most 14 vertices associated with a node, no node is assigned more than 14 incoming backbone links in step k, and Invariant 1 continues to hold after the node-allocation phase of step k.

Thus we have shown that both Invariants 1 and 2 continue to hold after step k.

Hence after step h_{max} (where $h_{max} \leq \frac{2}{\sqrt{a(n)}}$), each cell $\mathcal{D}$ has $O(\frac{h_{max}na(n)}{c}) \implies O(\frac{n\sqrt{a(n)}}{c})$ entering backbone links per channel, and each node appears on $O(h_{max}) = O(\frac{1}{\sqrt{a(n)}}) \implies O(\frac{n\sqrt{a(n)}}{c})$ (from Eqn. (10)) source backbones. ■

Detour backbones: From Lemma 21 the number of additional flows traversing a cell due to detour routing is only $O(\log^6 n)$, and each such flow will at most traverse the cell twice. Thus detour flows do not pose any significant load-balancing issue at any cell, and we can grow the backbones in $\mathcal{S}$ for these flows in any manner possible, i.e. by assigning links to any eligible node/channel (at least one eligible node is guaranteed to exist since, as a consequence of Lemma 17, each node can switch on at least $\lceil\frac{3f}{4}\rceil$ channels that are proper in the next cell).

Additional last hop: We now account for the possible additional last hop that some flows may have, yielding an additional cell in $\mathcal{S}$ (in addition to those traversed from source to pseudo-destination). We already argued that at most $O(na(n)) \implies O(\frac{n\sqrt{a(n)}}{c})$ flows (from Eqn. (9)) enter any cell on their additional last hop. Thus, even if their backbone links are assigned to the same channel/node, we would still have $O(\frac{n\sqrt{a(n)}}{c})$ flows per node and channel in any cell for the $\mathcal{S}$ stage.

b) Expanding backbone to $\mathcal{D}_b - \mathcal{S}$: In this stage $\mathcal{B}_p(x)$ expands into the cells traversed by flows for which x is the destination. Note that by our routing strategy a flow will only attempt to switch to the destination backbone when it enters *ready-for-transition* mode. From Lemma 23, the total number of flows traversing a cell in *ready-for-transition* mode is $O(\log^6 n)$ (counting possible repeat traversals), which is much smaller than $O(\frac{n\sqrt{a(n)}}{c})$. Thus flows on their destination backbone do not pose any major load-balance issues, and the backbones can be expanded into cells of $\mathcal{D}_b - \mathcal{S}$ by assigning links to any eligible node/channel.

B. Proving load-balance within a cell

We now show that no channel or interface bottlenecks form in the network when our described construction is used.

Per-Channel Load:

Lemma 25: The number of flows that enter any cell on a given channel is $O(\frac{n\sqrt{a(n)}}{c})$ w.h.p.

Proof: A flow on route $\mathcal{D}_1, \mathcal{D}_2, \ldots, \mathcal{D}_{j-1}, \mathcal{D}_j \ldots$ may enter a cell $\mathcal{D}_j$ on a channel i if (1) the flow is in *progress-on-source-backbone* mode, or it is in *ready-for-transition* mode, but is yet to find a transition into the destination

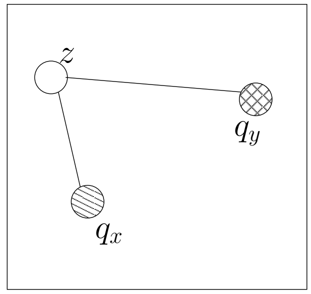

Fig. 5. Two additional transition links for a flow lying wholly within the cell

backbone , and i is the shared channel between the source backbone nodes in $\mathcal{D}_{j-1}$ and $\mathcal{D}_j$, or (2) the flow has already made a transition, and i is the shared channel between the destination backbone nodes in $\mathcal{D}_{j-1}$ and $\mathcal{D}_j$

We first consider the flows that enter a cell in *progress-on-source-backbone* mode, i.e., are proceeding on their source backbones. Recall that these are all non-detour-routed flows, since detour-routed flows are always in *ready-for-transition* mode. Then the number of such flows that enter any cell on a single channel is $O(\frac{n\sqrt{a(n)}}{c})$ from Lemma 24.

We now need to account for the fact that some of these flows may be in the *ready-for-transition* mode. From Lemma 23 there are $O(\log^6 n)$ flows traversing any cell in *ready-for-transition* mode w.h.p. (recall that these include the detour-routed flows with their repeat traversals counted separately, and the possible additional last $D'D$ hop). Thus regardless of whether they are still on their source backbone, or have already made the transition to their destination backbone, no channel can have more than $O(\log^6 n)$ such flows entering the cell.

Hence the number of flows entering on a single channel is $O(\frac{n\sqrt{a(n)}}{c}) + O(\log^6 n) \implies O(\frac{n\sqrt{a(n)}}{c})$ w.h.p. for each cell of the network. ■

Lemma 26: The number of flows that leave any cell on any single channel is $O(\frac{n\sqrt{a(n)}}{c})$ w.h.p.

Proof: Note that the flows that leave the cell, must then enter one of the 8 adjacent cells on that channel (as the corresponding backbone link for a flow leaves the current cell, and enters an adjacent cell). Thus, flows leaving the cell on a channel can be no more than 8 times the maximum number of flows entering a cell on any one channel, which has been established as $O(\frac{n\sqrt{a(n)}}{c})$ in Lemma 25. Hence, the total number of flows leaving any given cell on a given channel is also $O(\frac{n\sqrt{a(n)}}{c})$ w.h.p. ■

Lemma 27: The number of additional transition links scheduled on any single channel within any cell is $O(\log^6 n)$ w.h.p.

Proof: Recall that transition strategy outlined in the proof of Lemma 19, whereby the flow locates a cell along the route where the source backbone node q_x, and destination backbone node q_y are connected through a third node z. This yields two additional links $q_x \rightarrow z$, and $z \rightarrow q_y$ that lie entirely within the cell (Fig. 5). Note that the number of flows performing this transition in the cell can be no more than the number of flows traversing the cell in *ready-for-transition* mode. From Lemma 23 there are $O(\log^6 n)$ such flows traversing any cell w.h.p. In the worst case, we can count 2 additional links for each such flow as being all assigned to one channel. The result proceeds from this observation. ■

Per-Node Load:

Lemma 28: The number of flows that are assigned to any one node in any cell is $O(\frac{n\sqrt{a(n)}}{c})$ w.h.p.

Proof: A node is always assigned the single flow for which it is the source. A node is also assigned flows for which it is the destination, and from Lemma 7 there are at most $D(n) = O(\log n)$ such flows for any node w.h.p. Besides, a node may be assigned flows that are in the *ready-to-transition* mode, for which it facilitates a transition (if it is a *transition facilitator* node), or on whose destination backbone it figures. There are $O(\log^6 n)$ such transitioning flows in a cell w.h.p. from Lemma 23. Thus a node can only have $O(\log^6 n)$ such flows assigned.

We now consider the flows in *progress-on-source-backbone* mode that do not originate in the cell. These nodes are on their source-backbone, and from Lemma 24, each node has at most $O(\frac{n\sqrt{a(n)}}{c})$ such flows assigned. Thus, the resultant number of assigned flows per node is $1 + D(n) + O(\log^6 n) + O(\frac{n\sqrt{a(n)}}{c}) \implies O(\frac{n\sqrt{a(n)}}{c})$. ■

C. Transmission schedule

As mentioned earlier, from the Protocol Model assumption, each cell can face interference from at most a constant number β of nearby cells. Thus, if we consider the resultant cell-interference graph (a graph with a vertex for each cell, and an edge between two vertices if the corresponding cells can interfere with each other), it has a chromatic number at most $1+\beta$. Hence, we can come up with a global schedule having $1+\beta$ unit time slots in each round. In any slot, if a cell is active, then all interfering cells are inactive. The next issue is that of intra-cell scheduling. We need to schedule transmissions so as to ensure that at any time instant, there is at most one transmission on any given channel in the cell. Besides, we also need to ensure that no node is expected to transmit or receive more than one packet at any time instant.

We construct a conflict graph based on the nodes in the active cell, and its adjacent cells (note that the hop-sender of each flow shall lie in the active cell, and the hop-receiver shall lie in one of the adjacent cells, except for transition links, for which both lie in the active cell), as follows: we create a separate vertex for each flow for which a node in the cell needs to transmit data (we count repeat traversals or additional transition links as distinct flows for the purpose of scheduling; these have been accounted for while bounding the number of flows in a cell in previous lemmas). Since the flow has an assigned channel on which it operates in that particular hop, each vertex in the graph has an implicit associated channel. Besides, each vertex has an associated pair of nodes corresponding to the hop endpoints. Two vertices are connected by an edge if (1) they have the same associated channel, or (2) at least one of their associated nodes is the same. The scheduling problem thus reduces to obtaining a vertex-coloring of this graph. If we have a vertex coloring, then it ensures that (1) a node is never simultaneously sending/receiving for more than one flow (2) no two flows on the same channel are active simultaneously. Thus, the number of neighbors of a graph vertex is upper bounded by the number of flows requiring a transmission in the active cell on that channel, and the number of flows assigned to the flow's two hop endpoints (both hop-sender and hop-receiver). It can be seen from Lemma 25, Lemma 26, Lemma 27 and Lemma 28 that the degree of the conflict graph is $O(\frac{n\sqrt{a(n)}}{c}) + O(\frac{n\sqrt{a(n)}}{c}) + O(\log^6 n) + O(\frac{n\sqrt{a(n)}}{c}) + O(\frac{n\sqrt{a(n)}}{c}) = O(\frac{n\sqrt{a(n)}}{c})$ (note that $O(\log^6 n) \implies O(\frac{n\sqrt{a(n)}}{c})$, since we showed in Eqn. (9) that $\frac{n\sqrt{a(n)}}{c} = \Omega(\sqrt{\frac{n}{\log n}})$). Thus the graph can be colored in $O(\frac{n\sqrt{a(n)}}{c})$ colors.

Hence, the cell-slot (which can be assumed to be of unit time) is divided into $O(\frac{n\sqrt{a(n)}}{c}) = O(\frac{\sqrt{\frac{n\log n}{p_{rnd}}}}{c})$ equal length subslots, and all traversing flows get a slot for transmission. This implies that each flow gets a $\Omega(c\sqrt{\frac{p_{rnd}}{n\log n}})$

fraction of the time. Moreover, each cell gets at least one slot in $1+\beta$ slots, where β is a constant, and each channel has bandwidth $\frac{W}{c}$. Thus each flow gets a throughput of at least $\left(\frac{1}{1+\beta}\right)\left(\frac{W}{c}\right)\Omega(c\sqrt{\frac{p_{rnd}}{n\log n}}) = \Omega(W\sqrt{\frac{p_{rnd}}{n\log n}})$.

We thus obtain the following theorem:

Theorem 8: When $c = O(\log n)$ and $2 \leq f \leq c$, the per-flow network capacity with random (c, f) assignment is $\Theta(W\sqrt{\frac{p_{rnd}}{n\log n}})$.

VIII. A Remark on the Proof Technique

Note that many of our intermediate lemmas assume certain desirable events proved to occur w.h.p. in some of the lemmas proved before them, e.g., most intermediate lemmas are conditioned on the event in Lemma 11. It is not hard to see that the overall result continues to hold w.h.p., as briefly explained in this section:

Let a generic undesirable event be denoted by $\mathcal{E}_i$ (i.e., $\neg\mathcal{E}_i$ is the desirable event). We know from the union bound that:

$$Pr[\mathcal{E}_1 \cup \mathcal{E}_2] \leq Pr[\mathcal{E}_1] + Pr[\mathcal{E}_2] \tag{23}$$

Note that the following is also always true:

$$Pr[\mathcal{E}_1 \cup \mathcal{E}_2] = Pr[\mathcal{E}_1] + Pr[\neg\mathcal{E}_1]Pr[\mathcal{E}_2|\neg\mathcal{E}_1] \leq Pr[\mathcal{E}_1] + Pr[\mathcal{E}_2|\neg\mathcal{E}_1] \tag{24}$$

In light of this, it is not hard to see that the probability that even one of the undesirable events from any of these lemmas occurs, can be upper-bounded by summing up the individual (in some cases, conditional) probability of occurrence of each undesirable event, as bounded by each lemma (i.e., by essentially applying a union bound on the probabilities proved in each lemma). Since we have in all only a small constant number of lemmas, and each lemma shows that the (possibly conditional on events shown to occur w.h.p. in previous lemmas) probability of occurrence of some undesirable event goes to 0 (or equivalently shows that the probability of occurrence of the complementary desirable event goes to 1), the sum will also go to zero. Hence, the probability that even one of the undesirable events happens goes to 0.

IX. Discussion

In this paper, we described a construction that achieves a per-flow throughput of $\Omega(W\sqrt{\frac{p_{rnd}}{n\log n}})$ for $c = O(\log n)$, whenever c, f take values such that $f \geq 100$. For $f < 100$, the lower bound construction presented by us in [1], [9] (which yielded $\Omega(W\sqrt{\frac{f}{cn\log n}})$ per-flow throughput) is of the same asymptotic order (from Lemma 9, it follows that $\frac{\sqrt{\frac{f}{cn\log n}}}{\sqrt{\frac{p_{rnd}}{n\log n}}} = \Omega(\frac{1}{\sqrt{f}})$, and thus when $f < 100$, $\frac{\sqrt{\frac{f}{cn\log n}}}{\sqrt{\frac{p_{rnd}}{n\log n}}} = \Omega(1)$). In light of the upper bound of $O(W\sqrt{\frac{p_{rnd}}{n\log n}})$ proved by us in [1], [9], this establishes the capacity for random (c, f) assignment as $\Theta(W\sqrt{\frac{p_{rnd}}{n\log n}})$ in the regime $c = O(\log n)$.

We now show the following:

$$\begin{aligned} p_{rnd} &= 1 - (1 - \frac{f}{c})(1 - \frac{f}{c-1})...(1 - \frac{f}{c-f+1}) \\ &\geq 1 - (1 - \frac{f}{c})^f \geq 1 - e^{-\frac{f^2}{c}} (\text{ from Lemma 5}) \end{aligned} \tag{25}$$

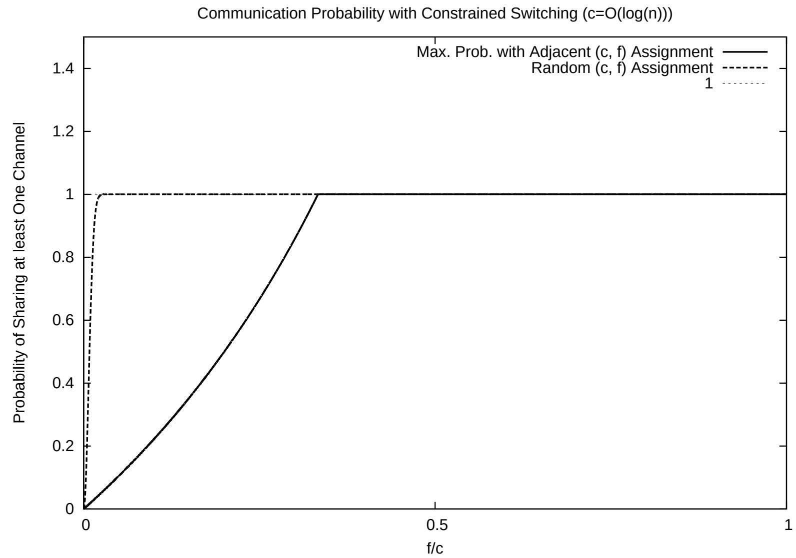

Fig. 6. Comparison of probability of sharing a channel

Thus, $f = \Omega(\sqrt{c}) \implies p_{rnd} = \Omega(1)$. To illustrate, if we set $f = \sqrt{c}$, $p_{rnd} \geq 1 - \frac{1}{e} > \frac{1}{2}$. In light of Eqn. (25), our result implies that $f = \Omega(\sqrt{c})$ suffices for achieving capacity of the same order as the unconstrained switching case [11]. For $f = \sqrt{c}$, the previously established lower bound of $\Omega(W\sqrt{\frac{f}{cn\log n}})$, would have yielded a capacity degradation of a factor of $c^{\frac{1}{4}}$, compared to the unconstrained switching case. In general, one may see that the capacity may diverge from the previous lower bound when $\frac{f}{c} \to 0$, but $f \to \infty$. Fig. 6 is a numerical plot (obtained by setting c to 10^4, and varying f from 2 to c) depicting how the probability p_{rnd} compares with the probability $p_{adj}^{max} = \min\{\frac{2f-1}{c-f+1}, 1\}$. Recall that p_{rnd} is the probability that two nodes share at least one channel in random (c, f) assignment, and p_{adj}^{max} is the upper bound on the probability that two nodes share at least one channel in adjacent (c, f) assignment [1]. It is quite remarkable that though both models allow nodes to switch between a subset of f channels, the additional degrees of freedom obtained via a random assignment lead to a much quicker convergence of p_{rnd} toward 1. The results in [1] established that connectivity was the dominant constraint determining capacity for adjacent (c, f) assignment in the $c = O(\log n)$ regime. The lower bound in this paper for random (c, f) assignment matches the upper bound imposed by the connectivity constraint (see [1]). Thus, the quick convergence of p_{rnd} to 1 leads to a quicker convergence of capacity towards that attainable via unconstrained switching.

It is to be noted that the lower bound of [1], [9] was obtained using a much simpler construction than the one described in this paper. Thus the two constructions represent an interesting trade-off in capacity versus scheduling/routing complexity.

X. Conclusion

We have presented a tight bound for capacity with random (c, f) assignment, for $c = O(\log n)$, $2 \leq f \leq c$. Our result indicates that capacity is $\Theta(W\sqrt{\frac{p_{rnd}}{n\log n}})$. Thus, one can achieve capacity of the same asymptotic order as unconstrained switching, when $f = \Omega(\sqrt{c})$. When $f < 100$, the capacity is achieved by using the construction for random (c, f) assignment described in [1], [9]. In this paper, we have described a new construction that achieves capacity for $f \geq 100$. We have also discussed the implications of this result, especially when compared to the capacity result for adjacent (c,f) assignment. There still remain some interesting open questions pertaining to the random (c,f) model, in terms of what is achievable via strictly asynchronous routing/scheduling. Other open issues include extension of the random and adjacent constraint models to multiple interfaces. Moreover, we believe that there is much potential for formalization and analysis of other kinds of switching constraint models.

References

[1] V. Bhandari and N. H. Vaidya, "Connectivity and Capacity of Multichannel Wireless Networks with Channel Switching Constraints," in *Proceedings of IEEE INFOCOM*, Anchorage, Alaska, May 2007.

[2] T. H. Cormen, C. E. Leiserson, and R. L. Rivest, *Introduction to Algorithms*. MIT Press, 1990.

[3] P. Gupta and P. R. Kumar, "The Capacity of Wireless Networks," *IEEE Transactions on Information Theory*, vol. IT-46, no. 2, pp. 388–404, March 2000.

[4] M. Mitzenmacher and E. Upfal, *Probability and computing*. Cambridge University Press, 2005.

[5] D. Dubhashi and D. Ranjan, "Balls and bins: a study in negative dependence," *Random Struct. Algorithms*, vol. 13, no. 2, pp. 99–124, 1998.

[6] A. E. Gamal, J. P. Mammen, B. Prabhakar, and D. Shah, "Throughput-Delay Trade-off in Wireless Networks." in *Proceedings of IEEE INFOCOM*, 2004.

[7] P. Hall, "On Representatives of Subsets," *J. London Math. Soc.*, vol. s1-10, no. 37, pp. 26–30, 1935. [Online]. Available: http://jlms.oxfordjournals.org

[8] H. Perfect, "The Mathematics of AGMs," *The Mathematical Gazette*, vol. 53, no. 383, pp. 13–19, Feb. 1969.

[9] V. Bhandari and N. H. Vaidya, "Connectivity and Capacity of Multichannel Wireless Networks with Channel Switching Constraints," Technical Report, CSL, UIUC, January 2007.

[10] P. Gupta and P. R. Kumar, "Critical power for asymptotic connectivity in wireless networks," in *Stochastic Analysis, Control, Optimization and Applications: A Volume in Honor of W.H. Fleming*, W. M. McEneany, G. Yin, and Q. Zhang, Eds. Boston: Birkhauser, 1998, pp. 547–566.

[11] P. Kyasanur and N. H. Vaidya, "Capacity of Multi-channel Wireless Networks: Impact of Number of Channels and Interfaces," in *MobiCom '05: Proceedings of the 11th annual international conference on Mobile computing and networking*. ACM Press, 2005, pp. 43–57.

[12] C. Fragouli, J. Widmer, and J.-Y. Le Boudec, "On the Benefits of Network Coding for Wireless Applications," in *4th International Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks*, 2006.

[13] Y. Azar, A. Z. Broder, A. R. Karlin, and E. Upfal, "Balanced Allocations," *SIAM J. Comput.*, vol. 29, no. 1, pp. 180–200, 2000.

[14] R. Motwani, R. Panigrahy, and Y. Xu, "Fractional Matching via Balls-and-Bins," in *RANDOM*, 2006.