Capacity Scaling of Wireless Networks with Inhomogeneous Node Density: Lower Bounds

Capacity Scaling of Wireless Networks With Inhomogeneous Node Density: Lower Bounds

Giusi Alfano, Michele Garetto, *Member, IEEE*, Emilio Leonardi, *Senior Member, IEEE*, and Valentina Martina

***Abstract*—We consider static ad hoc wireless networks comprising significant inhomogeneities in the node spatial distribution over the area and analyze the scaling laws of their transport capacity as the number of nodes increases. In particular, we consider nodes placed according to a shot-noise Cox process (SNCP), which allows to model the clustering behavior usually recognized in large-scale systems. For this class of networks, we propose novel scheduling and routing schemes that approach previously computed upper bounds to the per-flow throughput as the number of nodes tends to infinity.**

***Index Terms*—Ad hoc wireless networks, capacity, non-Poisson models, scaling laws.**

I. Introduction and Related Work

IN THEIR seminal work, Gupta and Kumar [1] evaluated the capacity of a static ad hoc wireless network consisting of n nodes randomly placed over a finite bidimensional domain and communicating among them (possibly in a multihop fashion) over point-to-point wireless links subject to mutual interference. They derived an upper bound[1] $O(1/\sqrt{n})$ to the per-node throughput, valid for arbitrary network topologies. In the case of nodes uniformly distributed over the network area, they proposed a scheme achieving $\Theta(1/\sqrt{n \log n})$ per-node throughput. Later on, Franceschetti *et al.* [2] have applied percolation theory results to show that $\Theta(1/\sqrt{n})$ transmission rate is achievable by the nodes also under uniform node placement.

The goal of our work is to extend the capacity scaling analysis to networks characterized by large inhomogeneities in the node density over the area. Indeed, almost all large-scale structures created by human or natural processes over geographical distances (such as urban or suburban settlements) exhibit significant degrees of clustering due to spontaneous aggregation of the nodes around a few attraction points. For example, many growing processes result into highly clustered networks due to preferential attachment phenomena. This fact motivated us to consider a general class of clustered point processes referred to as shot-noise Cox processes (SNCPs) [3], which includes several special cases widely used in different fields, such as Neyman–Scott process [4], Matérn cluster process [5], and Thomas process [6].

In this paper, we introduce and analyze a class of scheduling and routing schemes specifically tailored to clustered random networks, deriving constructive lower bounds to the per-node throughput as the number of nodes (and the number of clusters) tends to infinity. The obtained results get close (up to a poly-log factor) to the upper bounds derived in a separate paper [7] in all scenarios in which the system throughput is limited by interference among concurrent transmissions.

The main finding of this paper is that in order to approach the existing upper bound to the per-node throughput, it is necessary to employ scheduling and routing schemes that are significantly more involved than the ones proposed for networks with homogeneous node density [2].

To the best of our knowledge, only a few works have analyzed the capacity of clustered networks, especially departing from the assumption that nodes are uniformly placed over the network area. In [8], Toumpis considers a set of n nodes wishing to communicate to $m = \Theta(n^d)$ cluster heads and points out that the network throughput can be limited by the formation of bottlenecks at the clusters heads. Both sources and cluster heads are uniformly distributed, so the overall node density does not exhibit inhomogeneities.

The deterministic approach proposed in [9] allows to derive capacity results also for some nonuniform node distributions. In particular, the authors consider nodes distributed over $\sqrt{n}$ lines, or clustered into $\sqrt{n}$ neighborhoods. In both cases, a regular square tessellation of the network area can be built in such a way that no squarelet is empty w.h.p. (with high probability), while the maximum number of nodes in each squarelet increases at most as a poly-log function of n. Therefore, the network does not contain significant inhomogeneities, and the resulting capacity is similar to that derived by Gupta and Kumar.

In [10], the authors consider a system that contains many circular clusters with uniform node density within them, whose centers are distributed according to a Poisson process over the network area (a Matérn cluster process). Moreover, clusters are surrounded by a sea of nodes with much lower node density. The only quantity that scales with n is the network size: Below a critical network size, the per-node throughput is limited by the amount of data that a cluster can exchange with the sea of nodes, whereas above the critical size, the per-node throughput is limited by the capacity of the sea of nodes. In contrast to [10], we consider a more general SNCP, and we let the density of clusters

Manuscript received June 12, 2009; revised December 17, 2009; accepted April 13, 2010; approved by IEEE/ACM TRANSACTIONS ON NETWORKING Editor D. Rubenstein. Date of publication May 10, 2010; date of current version October 15, 2010. A preliminary version of this work has been presented at the IEEE INFOCOM 2009 conference.

G. Alfano, E. Leonardi, and V. Martina are with the Dipartimento di Elettronica, Politecnico di Torino, Torino 10129, Italy (e-mail: giuseppa.alfano@polito.it; emilio.leonardi@polito.it; valentina.martina@polito.it).

M. Garetto is with the Dipartimento di Informatica, Università di Torino, Torino 10124, Italy.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TNET.2010.2048719

[1]Given two functions $f(n) \geq 0$ and $g(n) \geq 0$: $f(n) = o(g(n))$ means $\lim_{n\to\infty} f(n)/g(n) = 0$; $f(n) = O(g(n))$ means $\limsup_{n\to\infty} f(n)/g(n) = c < \infty$; $f(n) = \omega(g(n))$ is equivalent to $g(n) = o(f(n))$; $f(n) = \Omega(g(n))$ is equivalent to $g(n) = O(f(n))$; $f(n) = \Theta(g(n))$ means $f(n) = O(g(n))$ and $g(n) = O(f(n))$; at last $f(n) \sim g(n)$ means $\lim_{n\to\infty} f(n)/g(n) = 1$.

(and the number of nodes per cluster) to scale with n as well. Moreover, our scheduling and routing schemes are different. In [11], the authors present a spatial framework to upper-bound the number of simultaneous transmissions in a network with general topology. However, their approach cannot be used to devise constructive schemes achieving the available per-node throughput for the class of networks considered here.

In [12], the authors consider a network in which nodes are arbitrarily placed on a domain of surface n under the constraint that the distance between any two nodes is larger than a given constant c. They derive both information theoretic upper bounds and constructive lower bounds employing advanced cooperative multiple-input–multiple-output (MIMO) transmission schemes. Our work is complementary with respect to [12] since our scheduling/routing strategies do not employ cooperative MIMO schemes. Furthermore, our model does not require any assumption on the minimum distance among nodes.

II. System Assumptions and Notation

A. Network Topology

We consider networks composed of a random number N of nodes (being $E[N] = n$) distributed over a square region $\mathcal{O}$ of edge length L, where L takes units of distance. The network physical extension L is allowed to scale with the average number of nodes since this is expected to occur in many growing systems. Throughout this work, we will always assume that $L = n^{\alpha}$, with $\alpha \geq 0$. To avoid border effects, we consider wraparound conditions at the network edges (i.e., the network area is assumed to be the surface of a bidimensional Torus).

The clustering behavior of large-scale systems is taken into account assuming that nodes are placed according to an SNCP. An SNCP over an area $\mathcal{O}$ can be conveniently described by the following construction. We first specify a point process $\mathcal{C}$ of cluster centers, whose positions are denoted by $\mathbf{C} = \{c_j\}_{j=1}^{M}$, where M is a random number with average $E[M] = m$. In the literature, the center points c_j are also called parent or mother points. Each center point c_j in turn independently generates a point process of nodes whose intensity at ξ is given by $q_j k(c_j, \xi)$, where $q_j \in (0, \infty)$ and $k(c_j, \cdot)$ is a dispersion density function, also called kernel, or shot. In the literature, the nodes generated by each center are also referred to as offspring or daughter points. The overall node process $\mathcal{N}$ is then given by the superposition of the individual processes generated by the cluster centers. The local intensity at $\xi \in \mathcal{O}$ of the resulting SNCP is

$$\Phi(\xi) = \sum_{j=1}^{M} q_j k(c_j, \xi).$$

Notice that $\Phi(\xi)$ is a random field in the sense that conditionally over all (q_j, c_j), the node process $\mathcal{N}$ is an (inhomogeneous) Poisson point process with intensity function Φ. We denote by $\mathbf{X} = \{X_i\}_{i=1}^{N}$ the collection of nodes positions in a given realization of the SNCP.

In this work, we restrict ourselves to kernels $k(c_j, \cdot)$ that are invariant under both translation and rotation, i.e., $k(c_j, \xi) = k(\|\xi - c_j\|)$ depends only on the euclidean distance $\|\xi - c_j\|$ of point ξ from the cluster center c_j. Moreover, we assume that $k(\|\xi - c_j\|)$ is a nonnegative, nonincreasing, bounded, and continuous function whose integral $\int_{\mathcal{O}} k(c_j, \xi) \mathrm{d}\xi$ over the entire network area is finite and equal to 1. In practice, the kernels considered in our work can be specified by first defining a nonnegative, nonincreasing continuous function $s(\rho)$ such that $\int_0^{\infty} \rho s(\rho) \mathrm{d}\rho < \infty$ and then normalizing it over the network area $\mathcal{O}$

$$k(c_j, \xi) = \frac{s(\|\xi - c_j\|)}{\int_{\mathcal{O}} s(\|\zeta - c_j\|) \mathrm{d}\zeta}.$$

Notice that in our asymptotic analysis we can neglect the normalizing factor $\int_{\mathcal{O}} s(\|\zeta - c_j\|) \mathrm{d}\zeta = \Theta(1)$. Indeed, $k(c_j, \xi) = \Theta(s(\|\xi - c_j\|))$. Notice that in order to have finite integral over increasing network areas, functions $s(\rho)$ must be $o(\rho^{-2})$, i.e., they must have a tail that decays with the distance faster than quadratically.

In the following, we will be especially interested in functions $s(\rho)$ whose tail decays as a power-law

$$s(\rho) = \min(1, \rho^{-\delta}), \quad \text{for } \delta > 2 \tag{1}$$

although our results apply to more general shapes as well.

Under the above assumptions on the kernel shape, quantity q_j equals the average number of nodes generated by cluster center c_j. We will assume for simplicity that all cluster centers generate on average the same number of nodes, hence $q_j = q$ for all $j = 1, \ldots, M$, although this restriction could be relaxed.

In our work, we let q scale with n as well (clusters are expected to grow in size as the number of nodes increases). This is achieved assuming that the average number of cluster centers scales as $m = n^{\nu}$, with $\nu \in (0, 1)$. Consequently, the number of nodes per cluster scales as $q = n^{1-\nu}$.

At last, we need to specify the point process $\mathcal{C}$ of cluster centers. We consider two different models.

- **Cluster Grid Model.** Clusters centers are deterministically placed over the vertices of a square grid.
- **Cluster Random Model.** Cluster centers are randomly placed according to a homogeneous Poisson process (HPP) of intensity $\phi_c = m/L^2$.

The Cluster Grid Model is simpler to analyze because the overall node process turns out to be a standard inhomogeneous Poisson point process whose intensity over the area can easily be evaluated since the locations $\mathbf{C}$ of clusters centers are assigned. This model serves as an intermediate step toward the analysis of the more complex Cluster Random Model.

For both models, we define

$$d_c = L/\sqrt{m} = n^{\alpha - \nu/2}. \tag{2}$$

This quantity represents, in the case of the Cluster Grid Model, the distance between two neighboring cluster centers on the grid; in the case of the Cluster Random Model, it is the edge of the square where the expected number of cluster centers falling in it equals 1. We call:

- *cluster-dense* regime the case $\alpha < \nu/2$, in which d_c tends to zero an n increases;
- *cluster-sparse* regime the case $\alpha > \nu/2$, in which d_c tends to infinity an n increases.

We leave for future studies the analysis of the case $\alpha = \nu/2$.

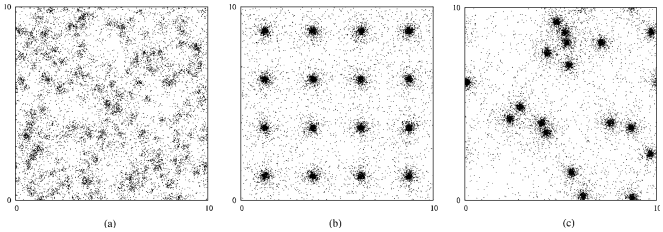

Fig. 1. Examples of topologies comprising $n = 10\,000$ nodes distributed over the square 10×10 ($\alpha = 0.25$). In all three cases, $s(\rho) \sim \rho^{-2.5}$. (a) Cluster Random Model with $\nu = 0.6$. This case belongs to the *cluster-dense* regime ($\alpha < \nu/2$). (b) Cluster Grid Model with $\nu = 0.3$. (c) Cluster Random Model with $\nu = 0.3$. Cases (b) and (c) belong to the *cluster-sparse* regime ($\alpha > \nu/2$).

Fig. 1 shows three examples of the kind of topologies considered in this paper, in the case of $n = 10\,000$ and $\alpha = 0.25$. In all three cases, we have assumed $s(\rho) = \min(1, \rho^{-2.5})$.

B. Communication Model

We assume that time is divided into slots of equal duration, and that in each slot the scheduling policy enables a set of transmitter–receiver pairs to communicate over point-to-point wireless links that are modeled as Gaussian channels of unit bandwidth. We consider point-to-point coding and decoding, hence signals received from nodes other than the (unique) transmitter are regarded as noise.

We assume that interference among simultaneous transmissions is described by the so called *generalized physical model*, according to which the rate achievable by node i transmitting to node j in a given time slot is limited to

$$R_{ij} = \log_2(1 + \mathrm{SINR}_{ij})$$

where SINR_{ij} is the signal-to-interference-plus-noise ratio at receiver j given that node i is transmitting to it.

We define Δ as the set of nodes that are enabled to transmit in a given slot, P_i as the power emitted by node i, ℓ_{ij} as the power attenuation between i and j, and N_0 as the ambient noise power. We have

$$\mathrm{SINR}_{ij} = \frac{P_i \ell_{ij}}{N_0 + \sum_{k \in \Delta, k \neq i,j} P_k \ell_{kj}}.$$

In this paper, we assume that all nodes employ the same power level while transmitting (i.e., $P_i = P, \forall i$). We remark that this assumption does not penalize the achievable throughput, as one can see by comparing our results to the upper bounds in [7], which are derived for arbitrary power assignment.

The power attenuation is assumed to be a deterministic function of the distance d_{ij} between i and j, according to $\ell_{ij} = d_{ij}^{-\gamma}$, with $\gamma > 2$. One drawback of this model is that the received power (and the corresponding rate) are amplified to unrealistic levels when d_{ij} tends to zero. Some authors have suggested to account for near-field propagation effect by bounding the attenuation function to 1 : $\ell_{ij} = \min\{1, d_{ij}^{-\gamma}\}$. However, any fixed bound leads to pathological throughput degradation in network regions where the node density tends to infinity, as pointed out in [14]. To avoid such problems, we simply assume that the achievable rate on any link cannot grow arbitrarily large, but is bounded by a constant R_0 due to the physical limitations of transmitters/receivers (maximum data speed of I/O devices, finite set of possible modulation schemes, etc.). Therefore, we consider the following variant of the *generalized physical model*:

$$R_{ij} = \min\{R_0, \log_2(1 + \mathrm{SINR}_{ij})\} \tag{3}$$

while keeping $\ell_{ij} = d_{ij}^{-\gamma}$, for any d_{ij}.

The proposed interference model satisfies the following basic property.

Lemma 1: Under the assumptions that, in a given time slot: 1) every transmission takes place between nodes whose relative distance is not greater than d; 2) the set Δ of transmitters is chosen by the scheduling policy in such a way that the distance $d_{tt'} = \|X_t - X_{t'}\|$ between any two transmitters t, t' in Δ is greater than or equal to $3d$; 3) nodes employ a common power level P; then the achievable rate by all concurrent transmissions according to (3) during the considered slot is

$$R(P, d) \geq \min\left(R_0, \log_2\left(1 + \frac{P d^{-\gamma}}{N_0 + cP d^{-\gamma}}\right)\right) \tag{4}$$

c being a positive constant.

Proof: We focus on a node r receiving data from a node t in the considered time slot, and we upper-bound the amount of interference I_r received by r due to concurrent transmissions all over the network.

By definition

$$I_r = P \int_0^\infty \ell(x) N_r(\mathrm{d}x) \tag{5}$$

where $\ell(x) = x^{-\gamma}$ is the power attenuation function at distance x, and $N_r(x)$ represents the number of transmitters (excluding t) within a distance x from r.

Now applying the triangular inequality, we can say that

$$N_r(x) \leq N_t(x + d) \tag{6}$$

where $N_t(x)$ denotes the number of interfering transmitters at a distance not greater than x from t.

Since the distance between any pair of transmitters is not smaller than $3d$, $N_t(x)$ can be easily bounded above by exploiting elementary geometric considerations. In particular, $N_t(x)$ is maximized when the transmitters are placed according to a regular triangular tessellation. In this case

$$N_t(x) = 0 \quad \text{for any } x < 3d \tag{7}$$

while

$$N_t(x) \leq \begin{cases} \frac{2\pi x^2}{9\sqrt{3}d^2}, & \text{if } 3d \leq x \leq L \\ \frac{2\pi L^2}{9\sqrt{3}d^2}, & \text{if } x > L. \end{cases} \tag{8}$$

Observe that from (6) and (7) it immediately follows that $N_r(x) = 0$ for $x < 2d$, while from (6) and (8) it follows $N_r(\infty) < \infty$. Hence, integrating (5) by parts, we obtain the following alternate expression for I_r:

$$I_r = -P \int_0^\infty N_r(x) \mathrm{d}\ell(x) = \gamma P \int_0^\infty x^{-(\gamma+1)} N_r(x) \mathrm{d}x \tag{9}$$

TABLE I
SYSTEM PARAMETERS

Symbol	Definition	scaling exponent
L	edge length of network area	$\alpha \geq 0$
m	average number of clusters	$0 < \nu < 1$
P	transmission power	≥ 0
δ	decay exponent of $s(\rho)$	n.a.
d_c	typical distance between cluster centres	$\alpha - \nu/2$
q	average number of nodes per cluster	$1 - \nu$

where, according to previous arguments, $l(0)N_r(0) = l(\infty)N_r(\infty) = 0$.

Hence, being $\gamma > 2$

$$I_r \overset{(9)}{=} \gamma P \int_0^\infty x^{-(\gamma+1)} N_r(x) \mathrm{d}x \tag{10}$$

$$\overset{(6),(8)}{\leq} \gamma P \int_{2d}^\infty x^{-(\gamma+1)} \frac{2\pi(x+d)^2}{9\sqrt{3}d^2} \mathrm{d}x = c\, P d^{-\gamma} \tag{11}$$

with $c = \frac{2\pi\gamma}{9\sqrt{3}2^{\gamma-2}}[\frac{1}{\gamma-2} + \frac{1}{\gamma-1} + \frac{1}{4\gamma}]$.

As a consequence, $d_{tr} \leq d$ being the distance between node t and node r, for the SINR at the receiver r, we obtain

$$\mathrm{SINR}_{tr} = \frac{P d_{tr}^{-\gamma}}{N_0 + I_r} \overset{(10)}{\geq} \frac{P d^{-\gamma}}{N_0 + c P d^{-\gamma}}.$$

Applying (3), we get the assertion. ∎

Furthermore, note that

$$\frac{P d^{-\gamma}}{N_0 + c P d^{-\gamma}} \geq \frac{P d^{-\gamma}}{2\max(N_0, c P d^{-\gamma})} = \min\left(\frac{1}{2c}, \frac{P d^{-\gamma}}{2N_0}\right)$$

and considering that for small values of x, $\log(1+x) \approx x$, we can claim that $R(P,d) = \Omega(\min[1, P d^{-\gamma}])$, where the product $P d^{-\gamma}$ scales with n as result of the particular parameters and architectural choices.

On this regard, note that for some combinations of parameters, the node density vanishes in the least populated regions of the network area; here, the typical distance between nodes tends to infinite, and transmissions between neighboring nodes are forced to cover increasingly large distances. As a consequence, the signal strength at the receiver, as well as the associated signal-to-noise ratio (SNR), vanishes unless nodes are able to increase their transmission power, so as to compensate for the signal power attenuation. Furthermore, observe that scaling P so as to overcompensate the link power attenuation (and thus produce an increasingly large SNR at the receiver) is not useful since, in our model, the rate of individual transmissions (3) is bounded by R_0.

In our analysis, we allow nodes to scale up their transmission power so as to achieve nonvanishing transmission rates. Actually, we do not make any specific assumption on how nodes select the transmission power P, thus obtaining results that are expressed parametrically as function of the particular choice of transmitted power P, which can depend arbitrarily on n as well.

Table I summarizes the parameters of our system. For the quantities that are allowed to scale with n, we have reported, in the third column, the restrictions on the scaling exponent in n, i.e., the assumptions on $\log_n(\langle \mathrm{parameter} \rangle)$. Note that d_c and q are not native parameters since they are derived from other parameters; however, we have included them in the table for convenience.

C. Traffic Model

Similarly to previous works [1], [2], we focus on *permutation traffic patterns*, i.e., traffic patterns according to which every node is source and destination of a single data flow at rate λ. Sources and destinations of data flows are randomly matched, establishing N end-to-end flows in the network. Note that a permutation traffic pattern is represented by a traffic matrix of the form $\mathbf{\Lambda} = \lambda \hat{\mathbf{\Lambda}}$, with $\hat{\mathbf{\Lambda}}$ being a permutation matrix (i.e., a binary-valued doubly stochastic matrix). Let $B(t)$ be the network backlog, that is, the number of data units already generated by sources that have not yet been delivered to destinations at time t. We say that traffic $\lambda \hat{\mathbf{\Lambda}}$ is *sustainable* if there exists a scheduling-routing policy such that $\limsup_{t\to\infty} B(t)/t = 0$ almost surely.

D. Asymptotic Analysis of Network Capacity

We are essentially interested in establishing how the network capacity scales with n under the assumptions we have introduced on network topology, communication model, and traffic pattern. To summarize, the quantities that depend on n are: 1) the network physical extension $L = n^\alpha$; 2) the number of cluster centers $m = n^\nu$, and consequently the average number $q = n^{1-\nu}$ of nodes belonging to the same cluster. As the number of nodes increases, we generate a sequence of systems indexed by n. The per-node throughput (or equivalently per-flow throughput) is $\Theta(h(n))$ if, given a sequence of random permutation traffic patterns with rate $\lambda^{(n)} = h(n)$, there exist two constants c, c' such that $c < c'$ and both the following properties hold:

$$\begin{cases} \lim_{n\to\infty} \Pr\{c\lambda^{(n)} \text{ is sustainable}\} = 1 \\ \lim_{n\to\infty} \Pr\{c'\lambda^{(n)} \text{ is sustainable}\} < 1. \end{cases}$$

Equivalently, we say in this case that the *network capacity* (or maximum network throughput) is $\Theta(nh(n))$. To simplify the notation, unless strictly necessary, we will omit the dependence of the variables on n in the following.

III. SUMMARY OF RESULTS

Table II summarizes the maximum achievable per-flow throughput (in order sense) under the *cluster dense* ($\alpha < \nu/2$) and *cluster sparse* ($\alpha > \nu/2$) regimes, for both Cluster Grid and Cluster Random models. The table reports the lower bounds (LB) obtained in this paper, together with the corresponding upper bounds (UB) derived in [7]. We observe that the lower bound coincides with the upper bound in the *cluster dense* regime, where the same per-node throughput $\Theta(1/\sqrt{n})$ as if nodes were uniformly placed over the domain can be achieved. In the *cluster sparse* regime, instead the per-node throughput drops below $1/\sqrt{n}$, for effect of inhomogeneities in the node spatial distribution, becoming tightly related to the minimum node density in the network.

To ease the comparison to the upper bounds, which are derived under the assumption that system performance is limited by the mutual interference among simultaneous transmissions (not by the transmission power), in Table III we have specialized the results of Table II for the *cluster-sparse* regime, when the transmission power P is selected by nodes in such a way to compensate the link power attenuation. This means that nodes select a transmission power $P = \Theta(\max(1, d_0^\gamma))$, where d_0 is

TABLE II
PER-FLOW THROUGHPUT ACHIEVABLE IN DIFFERENT CASES. LB (UB) STANDS FOR LOWER BOUND (UPPER BOUND)

	cluster-dense (LB=UB)	*cluster-sparse* (LB)	*cluster-sparse* (UB)
Cluster Grid	$\frac{1}{\sqrt{n}}$	$\max\left\{\frac{L\sqrt{qs(d_c)}}{n} R\left(P, \frac{1}{\sqrt{qs(d_c)}}\right), \frac{\sqrt{m}}{n} R(P, d_c)\right\}$	$\max\left\{\frac{L\sqrt{qs(d_c)}}{n}, \frac{\sqrt{m}}{n}\log n\right\}$
Cluster Random	$\frac{1}{\sqrt{n}}$	$\max\left\{\frac{L\sqrt{qs(d_c\sqrt{\log n})}}{n} R\left(P, \frac{1}{\sqrt{qs(d_c\sqrt{\log n})}}\right), \frac{\sqrt{m}}{n} R(P, d_c)\right\}$	$\max\left\{\frac{L\sqrt{qs(d_c)\log n}}{n}, \frac{\sqrt{m}}{n}\log n\right\}$

TABLE III
PER-FLOW THROUGHPUT ACHIEVABLE IN THE *CLUSTER-SPARSE* REGIME, WHEN $R(P,d) = \Theta(1)$, $\forall d$ AND $s(\rho) \sim \rho^{-\delta}$

	LB	UB
Cluster Grid	$\max\left\{\frac{L\sqrt{q}}{n\, d_c^{\delta/2}}, \frac{\sqrt{m}}{n}\right\}$	$\max\left\{\frac{L\sqrt{q}}{n\, d_c^{\delta/2}}, \frac{\sqrt{m}}{n}\log n\right\}$
Cluster Random	$\max\left\{\frac{L\sqrt{q}}{n\, d_c^{\delta/2}(\log n)^{\delta/2}}, \frac{\sqrt{m}}{n}\right\}$	$\max\left\{\frac{L\sqrt{q}}{n\, d_c^{\delta/2}(\log n)^{1/2}}, \frac{\sqrt{m}}{n}\log n\right\}$

the typical distance between transmitters and receivers in the least populated regions of the network area, so as to achieve a nonvanishing transmission rate on any link. Moreover, we have considered the most interesting case in which the node density function $s(\rho)$ decays as a power law with the distance from the cluster center, i.e., $s(\rho) \sim \rho^{-\delta}$ for $\rho \to \infty$. It can be easily seen in Table III that upper and lower bounds differ at most by a poly-log factor of n.

Results are graphically illustrated in Fig. 2, where we have reported, using a $\log_n$ vertical scale, the per-node capacity for fixed $\alpha = 0.3$, letting both δ and ν vary. Here, we have assumed that nodes adapt the transmission power to compensate the link power attenuation, leading to the results in Table III. Notice that on this scale, we can neglect poly-log factors in n, hence the represented surface corresponds to both upper and lower bounds.

The *cluster-dense* regime (i.e., $\alpha < \nu/2$), in which $\lambda = \Theta(1/\sqrt{n})$, produces the plateau at -0.5, for $\nu > 0.6$ any $\delta > 2$. In the *cluster-sparse* regime, the per-node capacity decreases for increasing values of δ and decreasing values of ν. The contour line at -0.7 splits the surface in two parts. The part below the contour line corresponds to the set of points (δ, ν) for which power adaptation is indeed necessary to have $R(P,d) = \Theta(1)$, $\forall d$, so that lower and upper bounds match (except for poly-log terms). For these values of (δ, ν), the node density within the least populated regions of the network tends to zero; thus, links established in these low-density regions have to cover increasingly large distances and require power adaptation to get a nonvanishing link rate. The part above the contour line, instead, does not require power adaptation, hence lower and upper bounds are tight even employing a constant power.

At last, in the lower part of the plot we can see that the surface becomes less steep. These portion of the surface corresponds to the set of points for which the $\max(\,)$ functions in Table III return the second term, hence the per-node capacity is $\lambda \sim \frac{\sqrt{m}}{n}$, with scaling exponent $\nu/2 - 1$ (independently of δ). In this case, as ν tends to 0, the per-node capacity approaches the lowest possible value $1/n$, with scaling exponent -1.

IV. PRELIMINARIES

Lemma 2: Consider a set of points $\mathbf{X}$ distributed over a bidimensional domain $\mathcal{O}$ of area L^2 according to a HPP of rate Φ. Let $\mathcal{T} = \{T_k\}$ be a regular tessellation of $\mathcal{O}$ (or any subregion of $\mathcal{O}$), whose tiles T_k have a surface $|T_k|$ not smaller than $16\frac{\log n}{\Phi}$.

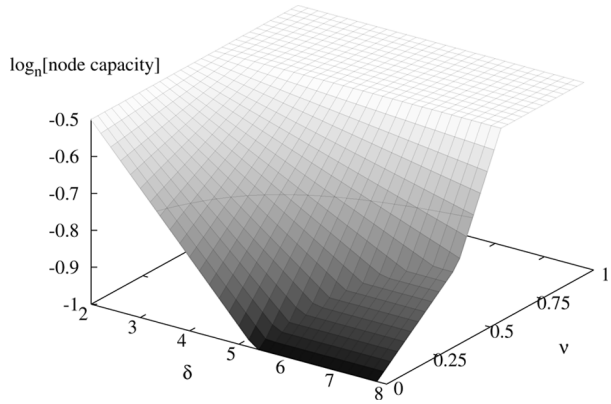

Fig. 2. Per-node capacity (in $\log_n$ scale) as a function of δ and ν, in the case $\alpha = 0.3$, assuming that power adaptation is applied.

Let $U(T_k)$ be the number of points of $\mathbf{X}$ falling within T_k. Then, uniformly over the tessellation, $U(T_k)$ is comprised between $\frac{\Phi|T_k|}{2}$ and $2\Phi|T_k|$, i.e., $\frac{\Phi|T_k|}{2} < \inf_k U(T_k) \le \sup_k U(T_k) < 2\Phi|T_k|$ with probability $p = 1 - O(\frac{1}{n})$.
The proof of this statement follows directly from the Chernoff bound and can be found in [2] and [15].

The previous result can be immediately generalized to the case of inhomogeneous Poisson processes (IPP).

Lemma 3: Consider a set of points $\mathbf{X}$ distributed over $\mathcal{O}$ according to an IPP having local intensity $\Phi(\xi)$, such that $\int_{\mathcal{O}} \Phi(\xi)\mathrm{d}\xi = n$. Let $\mathcal{T} = \{T_k\}$ be any tessellation of $\mathcal{O}$ (or any subregion of $\mathcal{O}$), whose tiles T_k satisfy: $\int_{T_k} \Phi(\xi)\mathrm{d}\xi \ge 16\log n$. Then, uniformly over the tessellations $1/2\int_{T_k}\Phi(\xi)\mathrm{d}\xi < \inf_k U(T_k) \le \sup_k U(T_k) < 2\int_{T_k}\Phi(\xi)\mathrm{d}\xi$, with probability $p = 1 - O(\frac{1}{n})$.

The following is a classical result on doubly stochastic matrices, known as the Birkhoff-von-Neumann (BvN) Theorem.

Lemma 4: (BvN Theorem). Any doubly stochastic matrix can be decomposed into a convex combination of permutation matrices.
From the BvN Theorem, it descends that every nonnegative, integer-valued matrix $A = [a_{ij}]$ can be decomposed into the sum of at most H subpermutation matrices (i.e., binary-valued doubly substochastic matrices), being $H = \max(\sum_i a_{ij}, \sum_j a_{ij})$ the maximum column/row sum [13].

At last, we report the main result of [2], together with a brief overview of their solution, which allows us to make a trivial generalization of the same result.

Lemma 5: Consider a set of nodes $\mathbf{X}$ placed according to a HPP with intensity $\Phi = 1$ over domain $\mathcal{O}$ of area $|\mathcal{O}| = n$; then, a scheduling/routing scheme exists such that the per-node throughput is $\Theta(1/\sqrt{n})$ under random permutation traffic with probability $p = 1 - O(\frac{1}{n})$.
The scheduling/routing policy proposed in [2] exploits the formation of several horizontal and vertical paths across the network in the transition region of an underlying percolation model. Nodes along these paths form a highway system that can carry information using multiple hops of length $\Theta(1)$ (recall that $|\mathcal{O}| = n$). The rest of the nodes access the highway system using single hops of length $O(\sqrt{\log n})$. The communication strategy is divided into four consecutive phases: In a first phase,

nodes drain their information to the highway; in a second phase, information is carried horizontally across the network through the highway; in a third phase, it is carried vertically; and in a last phase, information is delivered from the highway to the destination nodes. The system bottleneck turns out to be due to the phases in which information is carried over the highway, hence the per-node throughput is $\Theta(1/\sqrt{n})$.

The previous scheduling/routing strategy can be extended to the more general case in which nodes are placed over a domain $\mathcal{O}$ of area $|\mathcal{O}|$ according to a HPP of rate $\Phi = n/|\mathcal{O}|$. Employing the same communication strategy as in [2], all distances covered by transmissions are scaled by a factor $1/\sqrt{\Phi}$. When $\Phi = \Omega(1)$, the length of the hops along the highway system becomes $o(1)$ (the distance between transmitters and receivers on the highway is $\Theta(1/\sqrt{\Phi})$); the achievable rate on each hop along the highway remains $\Theta(1)$ (see Lemma 1), hence the per-node throughput is still $\Theta(1/\sqrt{n})$.

When $\Phi = o(1)$ instead, the distances covered by transmissions along the highway become $\omega(1)$, and the corresponding rate over each hop degrades to $R(P, 1/\sqrt{\Phi}) = o(1)$ for effect of the noise power at the receiver (see again Lemma 1), unless power levels are scaled up with n to perfectly compensate the link power attenuation. As a result, the per-node throughput becomes $\Theta(1/\sqrt{n})\ R(P, 1/\sqrt{\Phi}))$.

Thus, we obtain the following corollary.

Corollary 1: Consider a set of nodes $\mathbf{X}$ placed over domain $\mathcal{O}$ according to a HPP with intensity $\Phi = n/|\mathcal{O}|$. Then, a scheduling/routing scheme exists such that the per-node throughput is $\Theta(1/\sqrt{n})R(P, 1/\sqrt{\Phi}))$ under random permutation traffic with probability $p = 1 - O(\frac{1}{n})$.

Lemmas 2, 3, and 5 and Corollary 1 constitute the elementary building blocks to obtain our results. Thus, the majority of properties we are proving are verified with a probability $p = 1 - O(\frac{1}{n})$ when $n \to \infty$. In the following, we adopt the terminology "with high probability" (w.h.p.) to indicate events/properties that occur with a probability $p = 1 - O(\frac{1}{n})$; otherwise, we explicitly indicate the law according to which probability tends to 1, whenever necessary.

A. Asymptotic Analysis of the Local Intensity

Recall that under both the Cluster Grid and the Cluster Random models, the local intensity of nodes at point ξ can be written as $\Phi(\xi) = \sum_j qk(c_j, \xi)$. For both models, we define the following two quantities: $\bar{\Phi} = \sup_{\xi \in \mathcal{O}} \Phi(\xi)$ and $\underline{\Phi} = \inf_{\xi \in \mathcal{O}} \Phi(\xi)$.

The following theorem characterizes the asymptotic behavior of $\bar{\Phi}$ and $\underline{\Phi}$ in the Cluster Random Model.

Theorem 1: Consider nodes distributed according to a Cluster Random Model. Let $\eta(m) = d_c\sqrt{\log m}$. If $\eta(m) = o(1)$ (or equivalently $d_c = o(1/\sqrt{\log m})$), then it is possible to find two positive constants g_1, G_1 with $g_1 < G_1$ such that $\forall \xi \in \mathcal{O}$

$$g_1 \frac{n}{L^2} < \Phi(\xi) < G_1 \frac{n}{L^2} \quad \text{w.h.p.} \tag{12}$$

which means that $\underline{\Phi} = \Theta(\overline{\Phi})$. More in general, when $\eta(m) = \Omega(1)$, it is possible to find two positive constants g_2, G_2, such that, w.h.p., $\underline{\Phi} > g_2 q \log m\, s(d_c\sqrt{\log m})$ and $\bar{\Phi} < G_2 q \log m$.

Proof: The proof of this theorem, taken from [7], is reported in Appendix B for completeness. ■

Under the Cluster Grid Model, cluster centers are regularly placed in a deterministic fashion, hence $\bar{\Phi}$ and $\underline{\Phi}$ are two deterministic values depending only on system parameters.

Theorem 2: Consider nodes distributed according to a Cluster Grid Model. When $d_c = O(1)$, we have $\underline{\Phi} = \Theta(\bar{\Phi}) = \Theta(n/L^2)$. If $d_c = \omega(1)$, we have $\underline{\Phi} = o(\bar{\Phi})$, being $\underline{\Phi} = \Theta(q\, s(d_c))$ and $\bar{\Phi} = \Theta(q)$.

The proof of this theorem is a simplified version of the proof of Theorem 1 since, in the Cluster Grid Model, we do not have to deal with the randomness in the cluster centers positions. The proof can be obtained along the lines of Appendix B.

Recalling that the mean distance between cluster centers is $d_c = n^{\alpha-\nu/2}$ from (2), the above results show that, for both Cluster Grid and Cluster Random models, $\underline{\Phi} = \Theta(\bar{\Phi})$ in the *cluster-dense* regime, i.e., when $\alpha < \nu/2$ (this condition implies $d_c = o(1/\sqrt{\log m})$), whereas $\underline{\Phi} = o(\bar{\Phi})$ in the *cluster-sparse* regime, i.e., when $\alpha > \nu/2$ (this condition implies $d_c = \omega(1)$). In Section V, we analyze the case $\underline{\Phi} = \Theta(\bar{\Phi})$ under both Cluster Grid and Cluster Random models. In Section VI, we study the case $\underline{\Phi} = o(\bar{\Phi})$, separately considering the Cluster Grid and the Cluster Random models.

B. Thinning and Completing IPP and SNCP

Here, we introduce some useful techniques that can be applied to the point processes generated by our network model, which are essential for the construction of our scheduling-routing schemes.

We start with a couple of basic properties of Poisson point processes.

Lemma 6: Consider a set $\mathbf{X} = \{X\}_1^N$ of points distributed over a compact domain $\mathcal{O}$ according to an inhomogeneous Poisson process of intensity $\Phi(\xi)$. Let $\underline{\Phi}$ be $\inf_{\xi \in \mathcal{O}} \Phi(\xi)$. Then, for any $\Phi_0 \leq \underline{\Phi}$, it is possible to extract from $\mathbf{X}$ a subset of points $\mathbf{Z} \subseteq \mathbf{X}$ distributed over $\mathcal{O}$ according to a homogeneous Poisson process of rate Φ_0.

Proof: The proof is based on a standard thinning technique that is reported for completeness in Appendix A. ■

Lemma 7: Consider a set $\mathbf{X} = \{X\}_1^N$ of points distributed over a compact domain $\mathcal{O}$ according to an inhomogeneous Poisson process of intensity $\Phi(\xi)$. Let $\bar{\Phi}$ be $\sup_{\xi \in \mathcal{O}} \Phi(\xi)$. Then, it is possible to complete $\mathbf{X}$ (adding some extra points) to a superset $\mathbf{W} \supseteq \mathbf{X}$ of points distributed according to a homogeneous Poisson process of rate $\bar{\Phi}$.

Proof: The proof is reported for completeness in Appendix A. ■

The above two results can be immediately applied to the Cluster Grid Model since the overall point process generated by this model is a standard IPP. Therefore, we have the following.

Corollary 2: Consider nodes $\mathbf{X} = \{X\}_1^N$ with $\mathbb{E}[N] = n$, placed according to a Cluster Grid Model. Let $\bar{\Phi} = \sup_{\xi \in \mathcal{O}} \Phi(\xi)$ and $\underline{\Phi} = \inf_{\xi \in \mathcal{O}} \Phi(\xi)$ be the extreme values of the node intensity over the network domain $\mathcal{O}$, as computed in Theorem 2. Then, a subset of nodes $\mathbf{Z} \subseteq \mathbf{X}$ can be found with probability 1 such that $\mathbf{Z}$ forms a HPP with intensity $\underline{\Phi}$; moreover, $\mathbf{X}$ can be completed with probability 1 to a superset of points $\mathbf{W} \supseteq \mathbf{X}$ forming a HPP with intensity $\bar{\Phi}$.

Similar properties can be extended also to the Cluster Random Model, with the only difference that they hold w.h.p. (see

the last paragraph of Section IV), instead of with probability 1. This is stated more precisely in the following lemma.

Lemma 8: Consider nodes $\mathbf{X} = \{X\}_1^N$, with $\mathbb{E}[N] = n$, placed according to a Cluster Random Model. Then, a subset of nodes $\mathbf{Z} \subseteq \mathbf{X}$ can be found w.h.p. such that $\mathbf{Z}$ forms a HPP with intensity $\underline{\Phi}_0$, where $\underline{\Phi}_0 = g_1 \frac{n}{L^2}$ in the *cluster-dense* regime and $\underline{\Phi}_0 = g_2\, q \log ms (d_c \sqrt{\log m})$ in the *cluster-sparse* regime. Here, g_1 and g_2 are the constants defined in Theorem 1. Moreover, $\mathbf{X}$ can be completed w.h.p. to a superset $\mathbf{W} \supseteq \mathbf{X}$ that forms a HPP with intensity $\bar{\Phi}_0$, where $\bar{\Phi}_0 = G_1 \frac{n}{L^2}$ in the *cluster-dense* regime and $\bar{\Phi}_0 = G_2\, q \log m$ in the *cluster-sparse* regime. Again, G_1 and G_2 are the constants defined in Theorem 1.

Proof: We focus on the first property stated in the lemma (the thinning). The second property (the completion) follows exactly the same reasoning and is omitted for brevity. Recall that the Cluster Random Model is a doubly stochastic point process (a SNCP) with intensity $\Phi(\xi) = \sum_j q k(c_j, \xi)$, with $\mathbf{C} = \{c_j\}_1^M$ being the (random) collection of cluster centers' positions. By Theorem 1, $\Pr\{\underline{\Phi} \geq \underline{\Phi}_0\} = p_0 = 1 - O(\frac{1}{n})$ (both for the *cluster-dense* and *cluster-sparse* case). This means that, with probability p_0, the SNCP generates a topology in which $\underline{\Phi} = \inf_{\xi \in \mathcal{O}} \Phi(\xi) \geq \underline{\Phi}_0$. Let $\mathcal{C}_0$ be the set of cluster centers configurations for which $\underline{\Phi} \geq \underline{\Phi}_0$. By construction, the probability associated to $\mathcal{C}_0$ is p_0. For any cluster configuration $\mathbf{C} \in \mathcal{C}_0$, let $\underline{\Phi}_{|\mathbf{C}} = \inf_{\xi \in \mathcal{O}} \Phi(\xi \,|\, \mathbf{C})$ be the conditional minimum density of the point process. Conditionally over $\mathbf{C} = \{c_j\}_1^M$, the intensity $\Phi(\xi \,|\, \mathbf{C})$ is a deterministic function (i.e., conditionally over $\mathbf{C}$ the SNCP can be regarded as a standard IPP), hence we can apply Lemma 6 and extract a subset $\mathbf{Z}_{|\mathbf{C}}$ of nodes from $\mathbf{X}_{|\mathbf{C}}$ forming a HPP of rate $\underline{\Phi}_0$. Since this can be done for any $\mathbf{C} \in \mathcal{C}_0$, deconditioning over $\mathbf{C}$ we obtain that the thinning property holds with probability p_0, i.e., the probability of generating a cluster center configuration belonging to set $\mathcal{C}_0$. Since $p_0 = 1 - O(\frac{1}{n})$, the property holds w.h.p. for the Cluster Random Model. ∎

At last we report the following result that will come in handy in the next section.

Lemma 9: Consider nodes $\mathbf{X} = \{X\}_1^N$, with $\mathbb{E}[N] = n$, placed either according to a Cluster Grid or according to a Cluster Random Model in the *cluster-dense* regime (i.e., $\alpha < \nu/2$). Given a tessellation $\mathcal{T} = \{T_k\}$ of $\mathcal{O}$ with tiles T_k of area $|T_k| = \Omega(\frac{L^2 \log n}{n})$, let $U(T_k)$ be the number of points of $\mathbf{X}$ falling in tile T_k. Then, uniformly over the tessellation $U(T_k)$ satisfies w.h.p. $\frac{n}{L^2}|2T_k| \leq U(T_k) \leq \frac{n}{L^2}|2T_k|$.

Proof: Applying Corollary 2 for the Cluster Grid Model (Lemma 8 for the Cluster Random Model), we can extract from $\mathbf{X}$ w.p.1 (w.h.p.) a subset of points $\mathbf{Z}$ forming a HPP with intensity $\underline{\Phi}$ ($\underline{\Phi}_0 = g_1 \frac{n}{L^2}$), and we can complete $\mathbf{X}$ to $\mathbf{W}$ forming a HPP with intensity $\bar{\Phi}$ ($\bar{\Phi}_0 = G_1 \frac{n}{L^2}$). Then, setting $|T_k|$ not smaller than $16 \frac{\log n}{\underline{\Phi}}$ ($16 \frac{\log n}{\underline{\Phi}_0}$), the result immediately follows from Lemma 2, since by construction $\mathbf{Z} \subseteq \mathbf{X} \subseteq \mathbf{W}$.

V. Analysis of the *Cluster-Dense* Regime

The scheduling/routing strategies we propose are based on the idea of extracting a subset of nodes forming an infrastructure through which data can be transported across the network. This subset of nodes is distributed on the area according to an HPP. The rest of nodes communicate with the infrastructure using single-hop transmissions.

Theorem 3: In the *cluster-dense* regime, there exists a scheduling/routing scheme $\mathcal{P}_0$ achieving w.h.p. per-node throughput $\Theta(1/\sqrt{n})$.

Proof: In the Cluster Grid Model, we can apply Corollary 2 and extract with probability 1 a subset of nodes $\mathbf{Z} \subseteq \mathbf{X}$ distributed according to a HPP with intensity $\underline{\Phi} = \Theta(n/L^2)$; the same can be done, w.h.p., in the Cluster Random Model applying Lemma 8.

The proposed communication strategy is a generalization of the scheme introduced in [2]. Nodes in $\mathbf{Z}$ form the infrastructure carrying the traffic across the network area. Nodes in $\mathbf{Y} = \mathbf{X} \backslash \mathbf{Z}$ send/receive data from the infrastructure by single-hop communications with a close-by node in $\mathbf{Z}$.

Time is divided into regular frames, each one comprising three phases of equal duration. The first phase is exploited by nodes in $\mathbf{Y}$ to directly transmit to the nodes of $\mathbf{Z}$. The second phase is exploited to carry data through the infrastructure provided by nodes $\mathbf{Z}$; at last, the third phase is exploited to deliver data to $\mathbf{Y}$.

We start analyzing the first phase. First, we provide a construction that allows every node in $\mathbf{Y}$ to select a node of $\mathbf{Z}$ that acts as the first relay of its data stream. Then, we analyze the throughput that can be sustained during this first phase.

We partition $\mathcal{O}$ into squarelets of area $16 \log n/\underline{\Phi} \leq A \leq 32 \log n/\underline{\Phi}$ (the exact dimension of the squarelets is chosen is such a way to exactly cover $\mathcal{O}$ with an integer number of squarelets). According to Lemma 2, in every squarelet there are w.h.p. at least $\lfloor \underline{\Phi} A/2 \rfloor$ nodes belonging to $\mathbf{Z}$, while due to Lemma 9 the number of points of $\mathbf{X}$ in each squarelet is upper-bounded w.h.p. by $2A\bar{\Phi}$. Since, by construction, $\mathbf{Y} \subseteq \mathbf{X}$, any upper bound on the number of points $\mathbf{X}$ can be regarded as an upper bound on the number of points $\mathbf{Y}$.

Then, we uniformly partition the nodes of $\mathbf{Y}$ falling in each squarelet into $\lfloor \underline{\Phi} A/2 \rfloor$ groups, assigning each group to a different node of $\mathbf{Z}$ belonging to the same squarelet. By construction, each group contains at most $\lceil 4\bar{\Phi}/\underline{\Phi} \rceil = \Theta(1)$ nodes. Now, according to our scheme, only one transmission can take place in each squarelet at each time slot. Since both the transmitter and the receiver are placed in the same squarelet, their distance d_{tr} is upper-bounded by the diagonal of the squarelet: $d_{\mathrm{tr}} \leq \sqrt{2A}$, being $\sqrt{A} = \Theta(L\sqrt{\log n/n})$. Furthermore, to meet conditions of Lemma 1, transmissions occurring within the same squarelet, as well as transmissions occurring in squarelets containing points closer than $3\sqrt{2A}$, are to be orthogonalized over time. To do this, the squarelets are partitioned into a finite number of subsets, each subset comprising regularly spaced, weakly interfering squarelets. At any time, only one subset of squarelets is activated, and at most one transmission is enabled in each activated squarelet (this approach to limit interference among concurrent transmissions is pretty standard in related work; see, for example, [2, Theorem 3]). Applying Lemma 1, each enabled transmission can achieve a rate $\Omega(R(P, \sqrt{A}))$. However, we have to consider the fact that in each squarelet there are $O(\log n)$ competing nodes belonging to $\mathbf{Y}$. Since the fraction of time in which each of them can transmit scales as

$\Omega(1/\log n)$, the achievable average data rate sustainable by each node of $\mathbf{Y}$ during the first phase is

$$\lambda_1 = \Omega(R(P, L\sqrt{\log n/n})/\log n). \tag{13}$$

During the second phase, data are transported by the infrastructure provided by nodes $\mathbf{Z}$, either up to their final destination (if it belongs to $\mathbf{Z}$) or to the node $\mathbf{Z}$ assigned to the group of the destination. For the analysis of the second phase, we just apply the result of [2], adopting their scheduling/routing technique (notice that the scheme in [2] is itself divided into four subphases; see Lemma 1).

The only difference with respect to the assumptions of [2] relies on the fact in [2] a random permutation traffic pattern is assumed, in which each node is origin and destination of a single end-to-end flow established with a randomly selected destination. In our case instead, during the second phase, every node of $\mathbf{Z}$ pushes/pulls into/from the network infrastructure data belonging to (at most) $\lceil 4\bar{\Phi}/\underline{\Phi}\rceil + 1$ end-to-end flows. However, thanks to the BvN Theorem (see Lemma 1), we can decompose our traffic pattern into (at most) $\lceil 4\bar{\Phi}/\underline{\Phi}\rceil + 1$ permutation traffic patterns and devote to each of them a finite fraction of the system bandwidth.

Since $\mathbf{Z}$ is a HPP of intensity $\underline{\Phi}$ over an area L^2, the average number of nodes $E[|\mathbf{Z}|]$ composing the infrastructure equals $L^2\underline{\Phi}$. Since $\underline{\Phi} = \Theta(\frac{n}{L^2})$ w.h.p., we have $E[|\mathbf{Z}|] = \Theta(n)$. Applying Corollary 1, the average per-flow data rate sustainable in the second phase is, w.h.p.,

$$\lambda_2 = \Theta\left(\frac{1}{\sqrt{n}} R\left(P, \frac{L}{\sqrt{n}}\right)\right). \tag{14}$$

In the third phase, data are delivered from nodes in $\mathbf{Z}$ to their final destinations belonging to $\mathbf{Y}$ through single-hop transmissions. The analysis of this phase is identical to that of the first phase, exchanging the role of transmitters and receivers.

Since $\lambda_1 = \omega(\lambda_2)$, the second phase acts as the system bottleneck, and the per-node throughput is, w.h.p., $\lambda = \lambda_2 = \Theta(\frac{1}{\sqrt{n}} R(P, \frac{L}{\sqrt{n}}))$. At last, we observe that in the *cluster-dense* regime ($\alpha < \nu/2$), we always have $R(P, \frac{L}{\sqrt{n}}) = \Theta(1)$. This is because, being necessarily $\alpha < 1/2$ (since $\nu < 1$), the typical distance between transmitters and receivers during the second phase is $\frac{L}{\sqrt{n}} = O(1)$. We conclude that the per-node throughput is, w.h.p., $\lambda = \Theta(1/\sqrt{n})$. ■

VI. Analysis of the *Cluster-Sparse* Regime

Recall from Section IV-A that in the *cluster-sparse* regime ($\alpha > \nu/2$) we have $\underline{\Phi} = o(\bar{\Phi})$.

We distinguish two subcases, the *moderately sparse* regime when $\underline{\Phi} = \omega(1/d_c^2)$ and the *highly sparse* regime when $\underline{\Phi} = O(1/d_c^2)$, which are treated separately in Sections VI-A and B, respectively.

A. Moderately Sparse Regime

Our approach is still based on the idea of extracting a subset $\mathbf{Z}$ of nodes distributed according to a HPP of intensity $\underline{\Phi}$, forming the infrastructure carrying data across the network.

The per-node throughput guaranteed by such infrastructure is still given by Lemma 1. However, since $E[|\mathbf{Z}|] = L^2\underline{\Phi} = o(n)$, we expect a throughput degradation with respect to the case $\underline{\Phi} = \Theta(\bar{\Phi})$, in which $E[|\mathbf{Z}|] = \Theta(n)$.

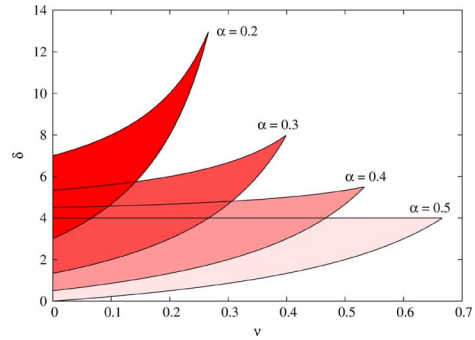

Fig. 3. Range of values of δ for which it is not convenient to employ direct transmissions to access the main transport infrastructure.

In principle, the rest of nodes could still access the above infrastructure through direct transmission, employing the same strategy $\mathcal{P}_0$ adopted in the *cluster-dense* regime. However, it turns out that this strategy is not always convenient and can actually result in severe throughput degradation. This is because, in highly populated areas, the density of nodes $\mathbf{Z}$ is negligible as compared to the density of nodes $\mathbf{X}$. Hence, the number of nodes $\mathbf{Y} = \mathbf{X} \setminus \mathbf{Z}$ that would compete for transmission to the same node of $\mathbf{Z}$ can be very large (the number of competing nodes is $\Theta(\bar{\Phi}/\underline{\Phi})$), shifting the system bottleneck to the access phase (phase 1).

This happens when λ_1, the per-flow average data rate achievable in phase 1, which would be given now by $\lambda_1 = R(P, \sqrt{\frac{\log n}{\underline{\Phi}}})\underline{\Phi}/\bar{\Phi}$, becomes negligible with respect to the per-flow average data rate achievable in phase 2, given by $\lambda_2 = \frac{L\sqrt{\underline{\Phi}}}{n} R(P, 1/\sqrt{\underline{\Phi}})$.

This condition, together with the assumption that $\underline{\Phi} = \omega(1/d_c^2)$, occurs when

$$2 + \frac{1-\nu}{\alpha - \frac{\nu}{2}} < \delta < \frac{1}{\alpha - \frac{\nu}{2}} - 2 \tag{15}$$

where we recall that δ is the decay exponent of the node intensity around a cluster center; see (1). Fig. 3 shows the range of values of δ as specified by (15), as a function of ν, for different values of α. Shaded areas correspond to those combinations of system parameters for which the system bottleneck would shift to the access phase if policy $\mathcal{P}_0$ were adopted.

Therefore, before accessing the infrastructure provided by $\mathbf{Z}$, data originated by densely populated area needs to be spread out evenly on the network area. This is accomplished gradually by covering highly dense areas of $\mathcal{O}$ with a hierarchy of intermediate, local transport infrastructures. Moving toward regions with higher and higher density, the area of such local infrastructures is reduced while keeping their transport capacity approximately the same. Intermediate infrastructures allow to reduce the distance covered by transmissions in highly populated regions while, at the same time, to efficiently balance the traffic toward the nodes of the main infrastructure.

We start describing our solution for the Cluster Grid Model, which is simpler to analyze, and later generalize our approach to the Cluster Random Model. For the sake of simplicity and

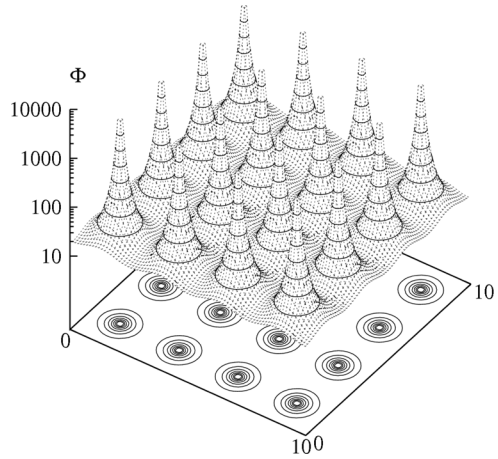

Fig. 4. Example of construction of nested domains $\mathcal{O}_k$ for the topology of the Cluster Grid Model depicted in Fig. 1(b).

TABLE IV
SUMMARY OF NOTATION

Symb.	Definition
$\mathcal{O}_k$	Layer-k infrastructure domain
$\mathcal{I}_k^j$	Layer-k isle around cluster centre c_j
A_k	Area of isles $\mathcal{I}_k^j$
$\mathbf{X}_k$	Nodes within $\mathcal{O}_k$
$\mathbf{Z}_k$	Nodes within $\mathcal{O}_k$ forming the layer-k infrastructure
$\mathbf{Y}_k$	Nodes within $\mathcal{O}_k \setminus \mathcal{O}_{k+1}$ not belonging to $\mathbf{Z}_k$
ζ_k	Density of nodes $\mathbf{Z}_k$ within $\mathcal{O}_k$

brevity, we assume that $\underline{\Phi} = \Omega(\log n/d_c^2)$, instead of $\underline{\Phi} = \omega(1/d_c^2)$. However, we remark that our results can be extended to the case $\underline{\Phi} = o(\log n/d_c^2)$ by slightly modifying the adopted scheduling policy.

1) Cluster Grid Model: We build a sequence of nested domains $\mathcal{O}_k$ (for $0 \leq k < K_{\max} = \lfloor \log_2(\bar{\Phi}/\underline{\Phi}) \rfloor$) according to the following definition:

$$\mathcal{O}_k = \{\xi \in \mathcal{O} : \Phi(\xi) \geq 2^k \underline{\Phi} = \zeta_k\} \tag{16}$$

i.e., $\mathcal{O}_k$ is the set of points in which the local intensity exceeds threshold ζ_k.

Note that, by construction, $\mathcal{O}_{k+1} \subset \mathcal{O}_k$, and $\mathcal{O}_0 = \mathcal{O}$. Moreover, while $\mathcal{O}_0 = \mathcal{O}$ is connected by hypothesis, domains $\mathcal{O}_k$ for $k \geq 1$ are in general disconnected, being composed of M congruent isolated domains $\mathcal{I}_k^j$ $(1 \leq j \leq M)$, hereinafter called isles, of area $|\mathcal{I}_k^j| = A_k$, surrounding each cluster center c_j:[2] $\mathcal{O}_k = \cup_j \mathcal{I}_k^j$. We call *pyramid* the set of isles $\mathcal{I}_k^j$ surrounding the same cluster center c_j (see Table IV). Fig. 4 shows an example of this construction for the topology in Fig. 1(b), characterized by $\underline{\Phi} = 19$, $\bar{\Phi} = 4754$, $K_{\max} = 7$. Notice that isles $\mathcal{I}_k^j$, $k \geq 1$ have approximately a circular shape. We define by $\mathbf{X}_k$ the restriction of $\mathbf{X}$ on $\mathcal{O}_k$; i.e., $\mathbf{X}_k$ comprises all points of $\mathbf{X}$ lying in $\mathcal{O}_k$. Our sequence of nested transport infrastructure can be obtained by virtue of the following result.

Proposition 1: For each k, a set of nodes $\mathbf{Z}_k \subseteq \mathbf{X}_k$ can be found such that: 1) $\mathbf{Z}_k$ is a HPP of intensity ζ_k on $\mathcal{O}_k$; 2) any node belonging to $\mathbf{Z}_{k-1}$ and to $\mathbf{X}_k$ also belongs to $\mathbf{Z}_k$, for $k \geq 1$.

Proof: We start from the top layer ($k = K_{\max}$) and apply the thinning strategy of Lemma 6 within domain $\mathcal{O}_{K_{\max}}$, extracting a set of nodes $\mathbf{Z}_{K_{\max}}$ distributed according to a HPP of intensity $\zeta_{K_{\max}}$. Next, we apply again the thinning procedure to set $\mathbf{Z}_{K_{\max}}$, extracting a set of nodes $\mathbf{Z}'_{K_{\max}-1}$ forming a HPP of intensity $\zeta_{K_{\max}-1}$ within $\mathcal{O}_{K_{\max}}$. We also independently extract a set of nodes $\mathbf{Z}''_{K_{\max}-1}$ forming a HPP of intensity $\zeta_{K_{\max}-1}$ within $\mathcal{O}_{K_{\max}-1} - \mathcal{O}_{K_{\max}}$. The nodes forming the infrastructure of layer $K_{\max} - 1$ are then obtained considering set $\mathbf{Z}_{K_{\max}-1} = \mathbf{Z}'_{K_{\max}-1} \cup \mathbf{Z}''_{K_{\max}-1}$. The above construction is repeated iteratively for each layer until we obtain the main infrastructure $\mathbf{Z}_0$. ■

Let $\mathbf{Y}_k = \mathbf{X}_k \setminus (\mathbf{X}_{k+1} \cup \mathbf{Z}_k)$, i.e., $\mathbf{Y}_k$ comprises those points of $\mathbf{X}_k$ lying in $\mathcal{O}_k - \mathcal{O}_{k+1}$ that do not belong to $\mathbf{Z}_k$. Table IV summarizes the notation.

Consider the following scheduling/routing scheme according to which time is partitioned into frames, each frame comprising $K_{\max}+1$ descending phases $K_{\max}, K_{\max}-1, \ldots, 1, 0$ followed by $K_{\max}+1$ ascending phases $0, 1, \ldots, K_{\max}$. Each phase is in turn partitioned into two periods.

Within descending phase k (here, index k runs from $K_{\max}$ down to 0), during the first period all nodes in $\mathbf{Y}_k$ are allowed to transmit their own data to a close node in $\mathbf{Z}_k$ within the same pyramid, equally distributing the traffic among the candidate receivers.

During the second period of the descending phase, every node in $\mathbf{Z}_k$ sends the data of which it is responsible to: 1) a randomly selected node lying within the same pyramid, and belonging to $\mathbf{Z}_k \cap \mathbf{Z}_{k-1}$, when $k > 0$; 2) a randomly selected nodes belonging to $\mathbf{Z}_0$, when $k = 0$. Again in this phase, source and destination nodes are matched in such a way to evenly distribute the traffic among the feasible destinations.

The data that are sent during the second period by every node of $\mathbf{Z}_k$ are those gathered in the previous phase (if $k < K_{\max}$), plus those gathered during the first period of the current phase, plus their own data. The data transport is achieved exploiting the scheme described in [2] on the infrastructure provided by $\mathbf{Z}_k$.

In ascending phase k, within the first period data directed to destinations in $\mathbf{X}_k \setminus \mathbf{X}_{k+1}$, for $k < K_{\max}$ (i.e., destinations lying in $\mathcal{O}_k$ but not in $\mathcal{O}_{k+1}$), or to destinations in $\mathbf{X}_{K_{\max}}$, for $k = K_{\max}$, are transmitted exploiting the scheme in [2] on the infrastructure provided by $\mathbf{Z}_k$, either directly to their destination (whenever the destination belongs to $\mathbf{Z}_k$) or to a close node in $\mathbf{Z}_k$, while at the same time, data directed to nodes lying within $\mathcal{O}_{k+1}$ (only for $k < K_{\max}$) are routed to nodes of $\mathbf{Z}_k \cap \mathbf{Z}_{k+1}$ lying within the same pyramid of the destination. During the second period of ascending phases $k \leq K_{\max}$, data directed to nodes in $\mathbf{Y}_k$ are delivered to their final destination through single-hop transmissions.

Fig. 5 illustrates an example of scheduling and routing of a flow established between two nodes belonging to different clusters (let these clusters be cluster 1 and cluster 2). More specifically, source node S belongs to set $\mathbf{Y}_3$ of cluster 1, whereas destination node D belongs to set $\mathbf{Y}_2$ of cluster 2. The figure shows one possible route for flow S–D, represented as a (logical) path connecting the source to the destination through a sequence of

[2]We can formally define domain $\mathcal{I}_k^j$, for any $k \geq 0$ and j, as the set of points of $\mathcal{O}_k$ whose closest cluster center is c_j; i.e., $\mathcal{I}_k^j = \{\xi \in \mathcal{O}_k : \|\xi - c_j\| \leq \|\xi - c_l\| \quad \forall l \neq j\}$. According to this formal definition, $A_0 = L^2/m$.

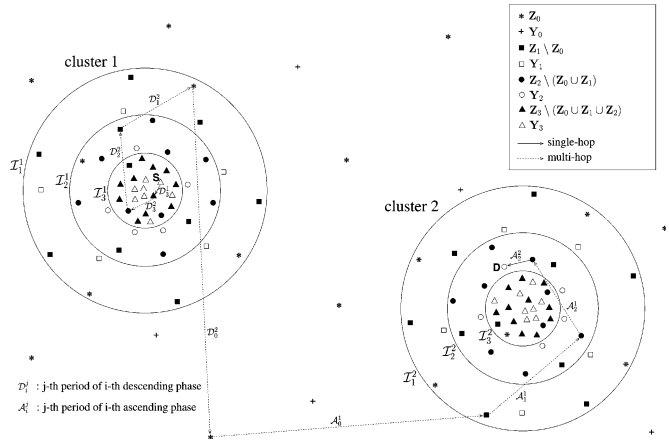

Fig. 5. Example of scheduling/routing of a flow established between node S belonging to cluster 1 and node D belonging to cluster 2. Nodes filled in white belong to sets $\mathbf{Y}_i$, whereas nodes filled in black belong to one or more of sets $\mathbf{Z}_i$. Different marks have been used to denote the nodes belonging to $\mathbf{Z}_i$ so as to illustrate one important property of our construction, namely the fact that if a node belongs to both $\mathbf{Z}_{k-1}$ and $\mathbf{X}_k$, it also belongs to $\mathbf{Z}_k$, for $k \geq 1$.

significant relay nodes. Each edge of the path can either be a single-hop transmission (solid line) or a multihop communication between the vertices (dashed line). Edges are also labeled with the phase and period in which the corresponding communication can be scheduled. The following notation has been used: $\mathcal{D}_i^j$ stands for the jth period of descending phase i; $\mathcal{A}_i^j$ stands for the jth period of ascending phase i. We observe that the chosen route is significantly more tortuous than other multihop routes between S and D that one could follow. This is due to the randomness introduced in our scheme in the selection of the next-hop relay belonging to a given set of nodes distributed around the cluster center. Nevertheless, the analysis below shows that such randomness does not penalize performance in order sense.

Theorem 4: Under the assumption that $\zeta_k A_k = \Omega(\beta^k \zeta_0 A_0)$, with A_k and ζ_k defined in (16), for some $\beta > 1$, the above described scheduling/routing scheme $\mathcal{P}_1$ for the Cluster Grid Model sustains a per-node throughput $\Theta(L\sqrt{\underline{\Phi}}/n \cdot R(P, d_0))$, being $d_0 = 1/\sqrt{\underline{\Phi}}$ with a probability $p = 1 - O(\frac{\log n}{n})$.

Proof: Consider descending phase k, with $0 \leq k \leq K_{\max}$. We fix the duration of this phase to $1/\sqrt{\beta^k}$ for some $\beta > 1$, and suppose that the two periods within the same phase are of equal duration. Notice that, being $\beta > 1$, the total duration of all of the descending phases is bounded, although their number tends to infinity. Fig. 6 illustrates the whole structure of the scheduling frame of our scheme, which is repeated indefinitely over time. We have also indicated in the figure the set of nodes that are allowed to transmit in each period of generic phase k (either descending or ascending).[3]

We start looking at the first period of this phase, in which nodes in $\mathbf{Y}_k$ are allowed to transmit to nodes in $\mathbf{Z}_k$ lying in the same pyramid. We apply Lemma 3 to the network region

[3]Notice that there is a slight asymmetry between descending and ascending phases since communications between the set of nodes $\mathbf{Z}_k$ and the set of nodes $(\mathbf{X}_k \setminus \mathbf{X}_{k+1}) \cap \mathbf{Z}_k$ have been assigned (for simplicity) to ascending phases only.

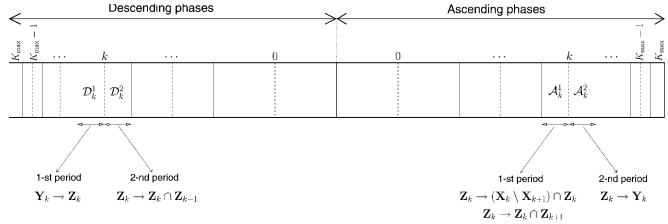

Fig. 6. Structure of the scheduling frame. In the bottom part of the figure, we have indicated the sets of nodes exchanging information in the two periods of generic phase k (either descending or ascending).

$\mathcal{O}_k \setminus \mathcal{O}_{k+1}$ [recalling that $\mathcal{O}_k$ is defined in (16)] if $k < K_{\max}$, and to $\mathcal{O}_{K_{\max}}$ otherwise, partitioning it into tiles T_k having surface[4] $|T_k| = \Theta(\log n/\zeta_k)$. Uniformly over the tiles, the number of points of $\mathbf{Y}_k$ falling in each tile is $O(\log n)$, whereas the number of points of $\mathbf{Z}_k$ in each tile is $\Omega(\log n)$ (actually, there are $\Theta(\log n)$ nodes of either sets in each tile). Hence, we can apply the same approach already employed in the case $\underline{\Phi} = \Theta(\bar{\Phi})$, assigning nodes $\mathbf{Z}_k$ to nodes $\mathbf{Y}_k$ residing within the same tile in such a way that the number of nodes $\mathbf{Y}_k$ associated to the same node $\mathbf{Z}_k$ is uniformly bounded by a constant. Also in this case, highly interfering transmissions are orthogonalized over time to meet conditions of Lemma 1: We adopt the same scheme used in the case $\underline{\Phi} = \Theta(\bar{\Phi})$, partitioning tiles into a finite number of subsets, each comprising mutually weakly interfering tiles. Since the number of conflicting transmitters per tile is $O(\log n)$, the fraction of time devoted to each transmitter is $\Omega(1/(2\sqrt{\beta^k}\log n)$, and consequently the average data rate achievable by every node in $\mathbf{Y}_k$ is

$$\lambda_1^k = \Omega(R(P, \sqrt{|T_k|})/(2\sqrt{\beta^k}\log n). \tag{17}$$

In the second period, for $k > 0$, data are transported within isles $\mathcal{I}_k^j$ by nodes $\mathbf{Z}_k$, adopting again the scheme of [2]. We observe that, by construction, in each isle $\mathcal{I}_k^j$ the amount of data to be transferred is that generated by nodes lying within the same isle. The number of nodes belonging to the infrastructure covering $\mathcal{I}_k^j$ is $z_k^j = \Theta(\zeta_k A_k)$ (note that we are assuming that $\zeta_k A_k = \Omega(\beta^k \zeta_0 A_0)$ being $\zeta_0 A_0 = \Omega(\log n)$). The total number of nodes n_k^j lying within $\mathcal{I}_k^j$ can be upper-bounded by $n_k^j = O(n/m)$ (note that $n/m = \omega(\log n)$, being $\nu < 1$).

Since the capacity of the infrastructure covering $\mathcal{I}_k^j$ is $\sqrt{z_k^j}R(P, d_k)$, being $d_k = 1/\sqrt{\zeta_k}$ (see Corollary 1), and considering that such infrastructure is used only for a fraction of time equal to $1/(2\sqrt{\beta^k})$, the aggregate data rate that can be sustained by the level-k infrastructure covering $\mathcal{I}_k^j$ is $\Omega(\frac{\sqrt{z_k^j}R(P,d_k)}{2\sqrt{\beta^k}}) = \Omega(\frac{\sqrt{\zeta_k A_k}R(P,d_k)}{2\sqrt{\beta^k}})$. Thus, every node of $\mathbf{Z}_k$ is allowed to exchange within level-k infrastructure a data rate that is $\Omega(\frac{\sqrt{\zeta_k A_k}R(P,d_k)}{2\sqrt{\beta^k} z_k^j})$. Considering that, by construction, every node of $\mathbf{Z}_k$ is pushing/pulling into level-k infrastructure the aggregate data of $O(n_z^j/z_k^j)$ end-to-end flows, the per-flow average data rate sustainable by level-k infrastructure is

$$\lambda_2^k = \Omega\left(\frac{\sqrt{\zeta_k A_k}R(P, 1/\sqrt{\zeta_k})}{2\sqrt{\beta^k} n_k^j}\right) \tag{18}$$

[4]We shape the tiles in such a way that the maximum distance between two points in the same tile T_k is $\Theta(\sqrt{|T_k|})$ for any k.

(to formally prove this result, we can resort to the trick of decomposing the traffic pattern into (sub)permutation traffic patterns).

In the case of the last descending phase $k = 0$, using similar arguments, it turns out that the per-flow average data rate sustainable by the ground infrastructure (level-0) is

$$\lambda_2^0 = \Theta(L\sqrt{\underline{\Phi}}/n)R(P, 1/\sqrt{\underline{\Phi}}).$$

Turning our attention to ascending phase k, we fix again the duration of this phase to $1/\sqrt{\beta^k}$ and further assume that the two periods within the same phase are of equal duration. Then, ascending phase k can be mapped to the corresponding descending phase k by reversing the time; i.e., by observing the data transmission process backward. As a consequence, the maximum throughput sustainable in ascending phase k equals the throughput sustainable in descending phase k.

We conclude that the maximum per-flow throughput sustainable by the whole system is given by the minimum among the per-flow throughput sustainable in every period of the frame. It turns out that the system bottleneck is due to the capacity of the ground infrastructure, i.e., $\lambda = \lambda_0^2 = \Theta(L\sqrt{\underline{\Phi}}/n)R(P, 1/\sqrt{\underline{\Phi}})$.

Observe that the assumption $\zeta_k A_k = \Omega(\beta^k \zeta_0 A_0)$ for some $\beta > 1$ is not restrictive. This property, indeed, can easily be verified to be equivalent to the condition that $s(\rho)$ decreases asymptotically to zero faster than $1/\rho^\delta$, for some $\delta > 2$.

As final remark, note that to obtain our result we need to jointly apply Lemmas 3 and 1 at every layer k of our construction. Since both results in Lemmas 3 and 1 hold with a probability $1 - O(\frac{1}{n})$, the whole construction will hold with a probability $p = 1 - O(\frac{K_{\max}}{n}) = 1 - O(\frac{\log n}{n})$. ■

2) Cluster Random Model: Recall that (Theorem 1), in the *cluster-sparse* regime of the Cluster Random Model, we have w.h.p. $\bar{\Phi} = O(q \log m)$ and $\underline{\Phi} = \Omega(q \log m\, s(d_c\sqrt{\log m}))$. Moreover, using exactly the same arguments of Theorem 1, we can strengthen the upper bound on the local density at point ξ, provided that we know the distance between point ξ and the closest cluster center, $d_{\min}(\xi) = \min_j \|\xi - c_j\|$. Then, $\Phi(\xi) = O(qs(d_{\min}(\xi)) \log n)$, w.h.p.

We employ a scheduling/routing scheme similar to the one devised for the Cluster Grid Model. The main difference lies in the fact that, in the Cluster Random Model, we have to deal with the irregular geometry induced by the random locations of the cluster centers.

We define domains $\mathcal{O}_k$, for $k \geq 1$, as follows: $\mathcal{O}_k = \{\xi \in \mathcal{O} : d_{\min}(\xi) \leq d_k = s^{-1}(\frac{2^{k-1}\zeta_1}{q})\}$, with $\zeta_1 = q\, s(\epsilon d_c)$, being ϵ a small positive constant (again, conventionally, $\mathcal{O}_0 = \mathcal{O}$). Note that the definition of $\mathcal{O}_k$ slightly differs from that introduced for the Cluster Grid Model.

Domains $\mathcal{O}_k$ are, in general, composed of many disjoint regions that are no longer associated through a one-to-one mapping to cluster centers c_j. Indeed, several cluster centers can now fall within the same connected region (see Fig. 7). Nevertheless, we will still denote by $\mathcal{I}_k^j$ the kth region to which cluster center c_j belongs to. Since domain $\mathcal{O}_k$ comprises all points whose distance from the closest cluster center is smaller than threshold d_k, regions $\mathcal{I}_k^j$ are associated to the connected components of the standard Gilbert's model of continuum percolation [16] with ball radius d_k.

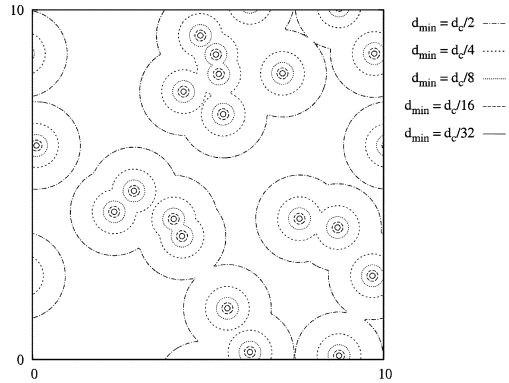

Fig. 7. Example of construction of nested domains for the topology of the Cluster Random Model depicted in Fig. 1(c).

The largest d_k is $d_1 = \epsilon d_c$. Choosing ϵ sufficiently small in such a way that the associated Gilbert's model is below the percolation threshold (we need $\epsilon < \epsilon^*$, where $\epsilon^* \approx 0.6$, [17]), we have the property that the maximum number of clusters centers belonging to the same region $\mathcal{I}_1^j$ is $O(\log n)$ w.h.p. [18]. Since by construction $\mathcal{O}_{k+1} \subset \mathcal{O}_k$, the same property holds for all $k > 1$. It follows that, in terms of physical extension, the area of region $\mathcal{I}_k^j$ lies w.h.p. in the interval $\pi(s^{-1}(\frac{2^{k-1}}{q}\zeta_1))^2 \leq |\mathcal{I}_k^j| \leq \pi(s^{-1}(\frac{2^{k-1}}{q}\zeta_1))^2 \log n$.

In addition, by construction the density of nodes within $\mathcal{O}_k$ is lower-bounded by $\zeta_k = 2^{k-1} q\, s(\epsilon d_c)$, for $k \geq 1$. Hence, it is possible to define for every $0 \leq k \leq K_{\max}$ (with $K_{\max} = \lfloor \log_2 \frac{q s(0)}{\zeta_1} \rfloor$) a set of points $\mathbf{Z}_k$ forming with probability 1 a HPP with intensity ζ_k on $\mathcal{O}_k$ (in the case of $k = 0$, $\mathcal{O}_0 = \mathcal{O}$ and $\zeta_0 = \underline{\Phi}$).

Then, we can apply the same scheme defined for the Clustered Grid Model. The restriction of $\mathbf{Z}_k$ to each $\mathcal{I}_k^j$ corresponds to a level-k regional infrastructure covering $\mathcal{I}_k^j$, whose aggregate capacity is $\sqrt{\zeta_k |\mathcal{I}_k^j|} R(P, 1/\sqrt{\zeta_k})$; such capacity is exploited during descending (ascending) phase k to transport the data originated from (destined to) nodes lying in $\mathcal{I}_k^j$.

Using the same arguments employed for the Cluster Grid Model, we can prove the following.

Theorem 5: Whenever $s(\rho)$ decreases faster than $1/\rho^\delta$, for some $\delta > 2$, the above-described scheduling/routing scheme $\mathcal{P}_2$ for the Cluster Random Model sustains a per-node throughput $\Theta(L\sqrt{\underline{\Phi}}/n \cdot R(P, d_0))$, with $d_0 = 1/\sqrt{\underline{\Phi}}$ with probability $p = 1 - O(\frac{\log n}{n})$.

B. Highly Sparse Regime

When the minimum node density $\underline{\Phi}$ drops below $1/d_c^2$ (in order sense), the approach of Section VI-A becomes suboptimal because the ground-level infrastructure provided by nodes $\mathbf{Z}_0$ (which determines the final network capacity) is too sparse (this especially occurs when $s(\rho)$ has a fast-decaying tail).

In this case, considering first the Cluster Random Model, a more dense set of nodes acting as ground-level infrastructure can be obtained by selecting one node per cluster—for example, according to an algorithm that selects for each cluster the closest node to the cluster center. It is of immediate verification that the obtained set of nodes $\mathbf{Z}_0'$ forms a HPP, in light of the fact that: 1) their number is distributed as a Poisson random variable with

average m; 2) the positions of points $\mathbf{Z}_0'$ are independent (since they belong to different clusters having independently located centers); 3) the marginal distribution of every point in $\mathbf{Z}_0'$ is uniform over $\mathcal{O}$ since cluster centers are uniformly distributed over $\mathcal{O}$. Note that, by construction, the density ζ_0' of $\mathbf{Z}_0'$ is: $\zeta_0' = m/L^2 = 1/d_c^2 = \omega(\underline{\Phi})$.

Then, defining domains $\mathcal{O}_k$ for $k \geq 1$ according to $\mathcal{O}_k = \{\xi \in \mathcal{O} : d_{\min}(\xi) \leq s^{-1}(\frac{2^k\zeta_0'}{q})\}$, a simple variant of the scheme proposed in Section VI-A2 permits to achieve higher throughput. We obtain the following result.

Theorem 6: When $\underline{\Phi} = O(1/d_c^2)$, a scheduling/routing scheme $\mathcal{P}_3$ can be defined that sustains a per-node throughput $\lambda = \Theta(\frac{L}{d_c n}R(P,d_c)) = \Theta(\frac{\sqrt{m}}{n}R(P,d_c))$.

Turning our attention to the Cluster Grid Model, by selecting one node per cluster according to the same algorithm proposed for the Cluster Random Model, it is still possible to obtain a set $\mathbf{Z}_0'$ of nodes providing a better ground-level infrastructure. In this case, the set $\mathbf{Z}_0'$ does not form a HPP. This fact, however, does not penalize the system performance since nodes in $\mathbf{Z}_0'$ are almost regularly spaced (indeed, using standard concentration results, it can be easily shown that, w.h.p., uniformly over the clusters, the distance between the selected node belonging to cluster j and the cluster center c_j is $o(d_c)$). As a consequence, standard results [1], [9] can be invoked to conclude that $\mathbf{Z}_0'$ provides an infrastructure of capacity $\Theta(\sqrt{m})$. Hence, we obtain the same per-node throughput as in Theorem 6.

VII. Conclusion

In this paper, we have derived constructive lower bounds to the asymptotic capacity of static ad hoc networks in which nodes are placed according to a shot-noise Cox process (SNCP). Such processes provide a fairly general model to capture the clustering behavior usually found in realistic large-scale systems. The presented lower bounds differ at most by a poly-log factor from existing upper bounds, under the assumption that the system throughput is limited by interference among concurrent transmissions. Our study has revealed the emergence of two regimes: the *cluster dense* regime, where an optimal $\Theta(1/\sqrt{n})$ per-node throughput is achieved, and the *cluster sparse* regime, where the per-node throughput degrades due to large inhomogeneities in the node spatial distribution. In the latter regime, we have shown that the system throughput is intrinsically related to the minimum node density within the network area.

Appendix A
Thinning and Completion of IPP

The thinning procedure works as follows: For any realization $\mathbf{X} = \{X\}_1^N$ of the IPP, consider each node and mark it with probability $g(\xi) = \frac{\Phi_0}{\Phi(\xi)}$, where ξ is the position of the node. The set of marked points form a Homogeneous Poisson Process with intensity Φ_0.

To show this fact, it is convenient to exploit a general result of the theory of point processes, stating that the distribution of a point process is completely specified by its void probabilities (see [19, Theorem 3.3]). Consider a domain $\mathcal{A} \in \mathcal{O}$. Let $U(\mathcal{A})$ be the random variable denoting the number of points of $\mathbf{X}$ falling within $\mathcal{A}$, and let $\hat{U}(\mathcal{A})$ be the random variable denoting the number of points of $\mathbf{Z}$ falling within $\mathcal{A}$. Since, given the event $\{U(\mathcal{A}) = i\}$, the i points of $\mathbf{X}$ falling in $\mathcal{A}$ are independently distributed over $\mathcal{A}$ according to the distribution $\frac{\Phi(x)}{\int_{\mathcal{A}}\Phi(x)\mathrm{d}x}$, it turns out that

$$\Pr\{\hat{U}(\mathcal{A}) = 0 \,|\, U(\mathcal{A}) = i\} = \left(1 - \frac{|A|Phi_0}{\int_{\mathcal{A}}\Phi(\xi)\mathrm{d}\xi}\right)^i.$$

Hence, unconditioning over i

$$\begin{aligned}\Pr\{\hat{U}(\mathcal{A}) = 0\} &= \sum_{i=0}^{\infty}\left(1 - \frac{|\mathcal{A}|\Phi_0}{\int_{\mathcal{A}}\Phi(\xi)\mathrm{d}\xi}\right)^i e^{(-\int_{\mathcal{A}}\Phi(\xi)\mathrm{d}\xi)}\frac{\left(\int_{\mathcal{A}}\Phi(\xi)\mathrm{d}\xi\right)^i}{i!}\\ &= e^{\int_{\mathcal{A}}\Phi(\xi)\mathrm{d}\xi - |\mathcal{A}|\Phi_0}\quad e^{-\int_{\mathcal{A}}\Phi(\xi)\mathrm{d}\xi} = e^{-|\mathcal{A}|\Phi_0}.\end{aligned}$$

The completion procedure is simply performed by adding to the existing points $\mathbf{X} = \{X\}_1^N$ some extra points independently distributed according to a IPP with intensity $\Phi_0 - \Phi(\xi)$, obtaining superset $\mathbf{W} \supseteq \mathbf{X}$.

Indeed, consider a domain $\mathcal{A} \in \mathcal{O}$. Let $U(\mathcal{A})$, $\tilde{U}(\mathcal{A})$, and $\bar{U}(\mathcal{A})$ be the number of points falling in $\mathcal{A}$ and belonging, respectively, to $\mathbf{X}$, $\mathbf{W} \setminus \mathbf{X}$, $\mathbf{W}$. It results

$$\begin{aligned}\Pr\{\bar{U}(\mathcal{A}) = 0\} &= \Pr\{U(\mathcal{A}) = 0\}\Pr\{\tilde{U}(\mathcal{A}) = 0\}\\ &= e^{-\int_{\mathcal{A}}\Phi(\xi)\mathrm{d}\xi}\, e^{-\Phi_0|\mathcal{A}| - \int_{\mathcal{A}}\Phi(\xi)\mathrm{d}\xi}\\ &= e^{-|\mathcal{A}|\Phi_0}.\end{aligned}$$

Appendix B
Proof of Lemma 2

Proof: The main steps of the proof are: 1) the domain $\mathcal{O}$ is divided into squarelets; 2) the local intensity at ξ_0 is expressed as sum of contributions, each due to cluster centers located in the same squarelet; 3) applying Lemma 2, every contribution is bounded w.h.p. (both from below and from above); 4) the upper (lower) bound is shown to converge w.h.p. to some value for $n \to \infty$.

In more detail, consider a generic point $\xi_0 \in \mathcal{O}$. By definition

$$\Phi(\xi_0) = \sum_{j=1}^{M} q\, k(\xi_0, c_j) = \sum_{j=1}^{M} q\frac{s(\|\xi_0 - c_j\|)}{\int_{\mathcal{O}} s(\|\zeta - c_j\|)\mathrm{d}\zeta}.$$

Now, let $\mathcal{A}$ denote a regular square tessellation of $\mathcal{O}$ such that each squarelet A_k has area $|A_k| = 16\,\eta^2(m)$, being $\eta(m) = d_c\sqrt{\log m}$. Let $\underline{d}_{0k}$ and $\bar{d}_{0k}$ be, respectively, the inferior and the superior of the distances between points $\xi \in A_k$ and ξ_0, i.e., $\underline{d}_{0k} = \inf_{\xi\in A_k}\|\xi - \xi_0\|$ and $\bar{d}_{0k} = \sup_{\xi\in A_k}\|\xi - \xi_0\|$; at last, let $\underline{U}(A_k)$ and $\bar{U}(A_k)$ be, respectively, a lower bound and an upper bound to the number of cluster centers falling in A_k. It results

$$\sum_k \frac{q}{H}s(\bar{d}_{0k})\underline{U}(A_k) < \Phi(\xi_0) < \sum_k \frac{q}{H}s(\underline{d}_{0k})\overline{U}(A_k)$$

being $H = \int_{\mathcal{O}} s(\|\zeta - c_j\|)\mathrm{d}\zeta$.

Applying Lemma 2, we have that, w.h.p., uniformly over k, $\underline{U}(A_k) \geq m/(2L^2)|A_k|$ and $\bar{U}(A_k) \leq 2m/L^2|A_k|$. Moreover, we observe that: 1) $\sum_k \frac{q}{H}s(\bar{d}_{0k})|A_k|$ and $\sum_k \frac{q}{H}s(\underline{d}_{0k})|A_k|$ can be interpreted, respectively, as lower Riemann sum and upper Riemann sum of $\int_{\mathcal{O}} \frac{q}{H}s(\|\xi - \xi_0\|)\mathrm{d}\xi$; 2) since $\eta(m) = o(1)$, the

mesh size of the partitions associated to Riemann sums vanishes to 0 as $n \to \infty$. As a consequence

$$\sum_k \frac{q}{H} s(\bar{d}_{0k})|A_k| \sim \sum_k \frac{q}{H} s(\underline{d}_{0k})|A_k| \sim \frac{q}{H} \int_{\mathcal{O}} s(\|\xi - \xi_0\|)\mathrm{d}\xi = q = \frac{n}{m} \quad (19)$$

and we conclude that

$$\frac{n}{2L^2} = q\frac{m}{2L^2} < \Phi(\xi_0) < q\frac{2m}{L^2} = \frac{2n}{L^2}.$$

Thus, (12) is verified for any $0 < g \leq 1/2$ and $G \geq 2$.

More in general, when $\eta(m) = \Omega(1)$, $\sum_k \frac{q}{H} s(\underline{d}_{0k})\bar{U}(A_k)$ and $\sum_k \frac{q}{H} s(\bar{d}_{0k})\underline{U}(A_k)$ provide, respectively, an upper bound and a lower bound to the local intensity. It turns out: $\sum_k \frac{q}{H} s(\underline{d}_{0k})\bar{U}(A_k) = \Theta(q \log m)$ and $\sum_k \frac{q}{H} s(\bar{d}_{0k})\underline{U}(A_k) = \Theta(q \log m\, s(d_c\sqrt{\log m}))$. ■

REFERENCES

[1] P. Gupta and P. R. Kumar, "The capacity of wireless networks," *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 388–404, Mar. 2000.

[2] M. Franceschetti, O. Dousse, D. N. C. Tse, and P. Thiran, "Closing the gap in the capacity of random wireless networks via percolation theory," *IEEE Trans. Inf. Theory*, vol. 53, no. 3, pp. 1009–1018, Mar. 2007.

[3] J. Møller, "Shot noise Cox processes," *Adv. Appl. Prob.*, vol. 35, pp. 614–640, 2003.

[4] J. Neyman and E. L. Scott, "Statistical approach to problems of cosmology," *J. Roy. Statist. Soc. B*, vol. 20, no. 1, pp. 1–43, 1958.

[5] B. Matérn, "Spatial variation," in *Lecture Notes in Statistics*, 2nd ed. Berlin, Germany: Springer, 1986, vol. 36.

[6] M. Thomas, "A generalization of Poissons binomial limit for use in ecology," *Biometrika*, vol. 36, no. 1/2, pp. 18–25, 1949.

[7] G. Alfano, M. Garetto, and E. Leonardi, "Capacity scaling of wireless networks with inhomogeneous node density: Upper bounds," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 7, pp. 1147–1157, Sep. 2009.

[8] S. Toumpis, "Capacity bounds for three classes of wireless networks: Asymmetric, cluster, and hybrid," in *Proc. ACM MobiHoc*, Tokyo, Japan, May 2004, pp. 133–144.

[9] S. R. Kulkarni and P. Viswanath, "A deterministic approach to throughput scaling in wireless networks," *IEEE Trans. Inf. Theory*, vol. 50, no. 6, pp. 1041–1049, Jun. 2004.

[10] E. Perevalov, R. S. Blum, and D. Safi, "Capacity of clustered ad hoc networks: How large is large?," *IEEE Trans. Commun.*, vol. 54, no. 9, pp. 1672–1681, Sep. 2006.

[11] A. Keshavarz-Haddad and R. H. Riedi, "Bounds for the capacity of wireless multihop networks imposed by topology and demand," in *Proc. ACM MobiHoc*, Montreal, QC, Canada, Sep. 2007, pp. 256–265.

[12] U. Niesen, P. Gupta, and D. Shah, "On capacity scaling in arbitrary wireless networks," *IEEE Trans. Inf. Theory*, vol. 55, no. 9, pp. 3959–3982, Sep. 2009.

[13] T. Weller and B. Hajek, "Scheduling nonuniform traffic in a packet-switching system with small propagation delay," *IEEE/ACM Trans. Netw.*, vol. 5, no. 6, pp. 813–823, Dec. 1997.

[14] O. Dousse and P. Thiran, "Connectivity vs capacity in dense ad hoc networks," in *Proc. IEEE INFOCOM*, Hong Kong, 2004, pp. 476–486.

[15] R. Motwani and P. Raghavan, *Randomized Algorithms*. Cambridge, U.K.: Cambridge Univ. Press, 1995.

[16] E. N. Gilbert, "Random plane networks," *J. SIAM*, vol. 9, pp. 533–543, 1961.

[17] J. Quintanilla, S. Torquato, and R. M. Ziff, "Efficient measurement of the percolation threshold for fully penetrable discs," *J. Phys. A, Math. Gen.*, vol. 33, pp. L399–L407, 2000.

[18] R. Meester and R. Roy, *Continuum Percolation*. Cambridge, U.K.: Cambridge Univ. Press, 1996.

[19] O. Kallenberg, *Random Measures*. Berlin, Germany: Akademie-Verlag, 1983.

Giusi Alfano was born in Naples, Italy, on March 22, 1978. She received the Laurea degree in communication engineering from the University of Naples Federico II, Naples, Italy, in 2001. From 2002 to 2004, she was involved in radar and satellite signal processing studies at the National Research Council and University of Naples. From April 2004 to October 2007, she was a Ph.D. student in information engineering at the University of Sannio, Benevento, Italy.

She currently holds a post-doctoral position with Politecnico di Torino, Torino, Italy. Her research work lies mainly in the field of random matrix theory applications to MIMO wireless communications and sensor networks and to the characterization of physical layers of random networks.

Michele Garetto (M'04) received the Dr.Ing. degree in telecommunication engineering and the Ph.D. degree in electronic and telecommunication engineering from Politecnico di Torino, Torino, Italy, in 2000 and 2004, respectively.

In 2002, he was a Visiting Scholar with the Networks Group of the University of Massachusetts, Amherst, and in 2004 he held a post-doctoral position with the Electrical and Computer Engineering Department of Rice University, Houston. He is currently an Assistant Professor with the University of Torino, Torino, Italy. His research interests are in the field of performance evaluation of wired and wireless communication networks.

Emilio Leonardi (M'99–SM'09) received the Dr.Ing. degree in electronics engineering and the Ph.D. degree in telecommunications engineering from Politecnico di Torino, Torino, Italy, in 1991 and 1995, respectively.

He is an Associate Professor with the Dipartimento di Elettronica, Politecnico di Torino. In 1995, he visited the Computer Science Department of the University of California, Los Angeles (UCLA). In the summer of 1999, he joined the High Speed Networks Research Group, Bell Laboratories/Lucent Technologies, Holmdel, NJ. In the summer of 2001, he was with the Electrical Engineering Department of Stanford University, Stanford, CA. His research interests are in the field of performance evaluation of wireless networks, P2P systems, and packet switching.

Valentina Martina was born in Torino, Italy, on August 3, 1982. She received the Laurea degree (summa cum laude) in mathematical modeling in engineering from Politecnico di Torino, Torino, Italy, in July 2007. Since January 2008, she has been pursuing the Ph.D. degree with the Dipartimento di Elettronica of Politecnico di Torino.