

# On the Data Gathering Capacity and Latency in Wireless Sensor Networks

Paolo Santi, *Member, IEEE*,  
IIT-CNR  
Pisa, ITALY

**Abstract**—In this paper, we investigate the fundamental properties of data gathering in wireless sensor networks, in terms of both capacity and latency. We consider a scenario in which  $s(n)$  out of  $n$  total network nodes have to deliver data to a set of  $d(n)$  sink nodes at a constant rate  $\lambda(n, s(n), d(n))$ . The goal is to characterize the maximum achievable rate, and the latency in data delivery. We present a simple data gathering scheme that achieves asymptotically optimal data gathering capacity and latency with arbitrary network deployments when  $d(n) = 1$ , and for most scaling regimes of  $s(n)$  and  $d(n)$  when  $d(n) > 1$  in case of square grid and random node deployments. Differently from most previous work, our results and the presented data gathering scheme do not sacrifice energy efficiency to the need of maximizing capacity and minimizing latency. Finally, we consider the effects of a simple form of data aggregation on data gathering performance, and show that capacity can be increased by a factor  $f(n)$  with respect to the case of no data aggregation, where  $f(n)$  is the node density. To the best of our knowledge, the ones presented in this paper are the first results showing that asymptotically optimal data gathering capacity and latency can be achieved in arbitrary networks in an energy efficient way.

*Index terms:* Wireless sensor networks, data gathering capacity, data gathering latency, physical interference model.

## I. INTRODUCTION

After several years of intensive research and technological development, wireless sensor networks are becoming increasingly deployed in a wide range of application scenarios, and will likely be part of our everyday life in a few years. Despite that, our understanding of fundamental properties of these networks – the so-called *scaling laws* –, which can help the network designer discriminating what can and cannot be done in such networks, is still limited. In particular, *network capacity* and *communication latency* are two fundamental properties of a wireless sensor network: the former can help characterizing the overall amount of data that can be transported within the network, while the latter can help understanding, for instance, whether the network can be effectively used for real-time event detection. When a wireless sensor network is to be used for, e.g., environmental monitoring, in which an extensive set of sensor readings has to be collected at one or few data collection sites (sinks), the combined characterization of network capacity and communication latency scaling laws is fundamental to understanding feasibility of a large scale network deployment.

Starting from the seminal Gupta and Kumar’s work on capacity [7] for unicast transmissions, fundamental properties of general wireless multihop networks have been extensively studied in the literature in the last few years (see Section II

for details). These studies have shown a considerable influence of the network topology and traffic patterns on the achievable network capacity and latency limits. For instance, in [7] it has been shown that while the per-node capacity in a network composed of  $n$  nodes degrades as  $O\left(\frac{1}{\sqrt{n}}\right)$  when  $n$  grows to infinity in arbitrary networks, it degrades at a faster pace (as  $O\left(\frac{1}{\sqrt{n \log n}}\right)$ ) if nodes are randomly deployed. The above results hold under the assumption that traffic source/destination pairs are randomly chosen. However, if communications are localized (i.e., the destination of a packet is randomly chosen amongst one of the source’s immediate neighbors), the situation is drastically different, and per-node throughput actually remains *constant* as  $n$  grows to infinity. Another example of the influence of the traffic pattern on network capacity and communication latency can be derived by comparing the findings of [13], which show that asymptotically optimal capacity and latency cannot be simultaneously achieved for unicast transmissions, with those of [15], which show that asymptotically optimal capacity and latency can indeed be simultaneously achieved for broadcast communications.

Given the above discussion, it is clear that characterizing the capacity and latency scaling laws of a wireless sensor network requires taking into account the typical traffic patterns of this type of networks. In this paper, we consider one such typical pattern – the *data gathering* pattern –, in which many (or all) the sensor nodes have to regularly deliver data to one (or few) sink node(s). The data gathering network capacity has been recently investigated in a few papers [4], [11], [12], [24] (see Section II for details), which, however, are based on quite strict assumptions for what concerns network topology, properties of the observed random field, and/or PHY layer capabilities. When simultaneous transmissions are allowed in the network [11], [12], radio interference is modeled using a relatively inaccurate interference model (the protocol interference model [7]). Furthermore, the effects of data aggregation strategies on the achievable network capacity have only been partially considered. Most importantly, none of the existing works is concerned with *energy efficiency* of the communications, nor considers *latency* in data communication. There is no need to comment on the importance of energy efficient communication in wireless sensor networks. As for latency, we stress that latency is a fundamental property in many wireless sensor network application scenarios, for instance when the network is used for real-time event detection.

The main findings of this paper, which are extensively

described at the end of Section II, are that asymptotically optimal data gathering capacity *and* latency can be simultaneously achieved in arbitrary and random networks in an energy efficient manner, as long as the network deployment satisfies a property which we call *minimal cell connectivity*. The data gathering capacity is limited by the number of sink nodes in the network, and can thus be improved (up to a certain point) by increasing the number of data collection sites. Another way of improving the data gathering capacity is by exploiting a very simple form of data aggregation, which increases network capacity of an  $O(f(n))$  factor, where  $f(n)$  is the node density.

The rest of this paper is organized as follows. In Section II, we discuss related work and extensively describe this paper's contributions to the state-of-the-art. In Section III, we introduce the network model used in the rest of the paper. In Section IV, we consider the case of single sink networks, and present a data gathering scheme that provides asymptotically optimal data gathering capacity and latency. In Section V, we extend the data gathering scheme to the case of multiple sink networks, and characterize conditions under which asymptotically optimal data gathering capacity can be achieved. Section VI discusses energy efficiency of the data gathering schemes proposed in sections IV and V. Finally, Section VII summarizes the main results of the paper and discuss their implications for the design of wireless sensor network data gathering applications.

## II. RELATED WORK AND CONTRIBUTIONS

The investigation of fundamental wireless network properties has been subject of intensive research since the seminal Gupta and Kumar's work [7]. In [7], the authors lay down the foundations of wireless multihop networking, showing that short-range communications have to be preferred to long-range communications if the objective is to maximize network capacity. This is due to the fact that short-range communications minimize the interference level in the network, allowing for a degree of spatial reuse (i.e., simultaneous transmissions) that is sufficient to outweigh the relay burden caused by multihop communications. However, this relay burden considerably limits per-node network capacity, which is shown to converge to 0 as the number  $n$  of network nodes grows to infinity. In [6], Grossglauser and Tse show that the Gupta and Kumar's capacity limit does not apply in mobile networks: the main idea is to exploit node mobility, rather than multihop forwarding, to deliver packets to the destinations. This way, per-node capacity can be optimized (i.e.,  $\Omega(W)$  per-node capacity can be achieved, where  $W$  is the channel capacity), at the price of considerably increasing communication latency. The tradeoff between network capacity and latency in packet delivery is investigated, e.g., in [13], [20]. In particular, in [13] the authors show that the achievable per-node capacity is upper bounded by  $O(\frac{D}{n})$ , where  $D$  is the average packet delay. This result implies that  $\Omega(W)$  per-node capacity can be achieved only by allowing  $O(n)$  latency, which is not optimal in most two-dimensional deployment scenarios.

The above results refer to the unicast communication pattern. Other traffic patterns have been recently investigated, such as multicast [10], [19] and broadcast [8], [9], [15], [16],

[17], [23]. We briefly comment on the results concerning broadcast communications, which are indeed relevant to wireless sensor networks (think about a sink node broadcasting a query in the network, downloading new software to the nodes, etc.). In [23], Zheng investigates the broadcast capacity of random networks with single broadcast source, and presents a broadcast scheme providing asymptotically optimal capacity. The author also studies the information diffusion rate, which is closely related to latency. However, broadcast capacity and information diffusion rate are not jointly considered, i.e., the capacity-achieving broadcast scheme does not optimize diffusion rate, and vice-versa. In [8], [9], the authors consider a more general network model, in which arbitrary node positions and arbitrary number of broadcast sources are allowed. The results of [8], [9] essentially confirm the finding on broadcast capacity reported in [23] in the more general model. In [15], the authors consider a network model similar to that of [9] (but assuming a single broadcast source), and show that asymptotically optimal broadcast capacity and latency can be achieved, subject to a (not very restrictive) condition on network topology. This result has been recently generalized to the case of arbitrary number of sources in [16], and to the case of mobile nodes [17].

The data gathering capacity of wireless networks has been investigated in a few recent papers. In [12], the authors consider a model in which nodes are evenly spaced in a circular area. Nodes monitor a random, spatially-correlated field, and the goal is to deliver snapshots of this random field to a sink, located in the center of the area. As the node density increases, correlation of the sensed data increases, as well as the demand of traffic directed to the sink. Since correlated data can be compressed, the question addressed in [12] is whether data compression can be exploited to deliver accurate snapshots of the random field (with an upper bounded MSE) to the sink. The authors essentially prove a negative result, showing that the benefits of data compression are not sufficient to outweigh the capacity bottleneck at the sink node, which implies that the per-source available capacity is  $O(\frac{W}{n})$ . In [4], El Gamal shows how to go beyond this per-node capacity limit achieving  $O(\frac{W \log n}{n})$  per-source capacity, which is proved to be sufficient to guarantee delivery of accurate snapshots of a Gaussian, bandwidth-limited random field. The main idea to improve capacity is exploiting advanced PHY layer techniques, namely cooperative communication: every node distributes its sampled value to neighboring nodes, and then all these nodes cooperatively transmit the information to the sink node, realizing distributed beamforming. Thus, owing to the beamforming gain, capacity at the sink node can be increased of a logarithmic factor. Although interesting from a theoretical point of view, El Gamal's result is hard to apply in a practical scenario, due to the many assumptions on which it is based: nodes are evenly spaced on a sphere, with the sink node located in the center of the sphere; at the PHY layer, nodes are assumed to perform transmit power control, to know the radio signal phase shifts between all possible pairs of nodes, to be able to perform cooperative communication, etc.; finally, the joint distribution of the random field observed

at the various nodes is assumed to be known to all sensor nodes. More recently, Zheng and Barton [24] have shown the benefits of cooperative communication in a less constrained environment, where nodes are randomly distributed, use a common transmit power, and a-priori statistical knowledge of the observed random field and of the channel characteristics are not needed. In particular, the authors show that even in this less constrained environment  $O\left(\frac{W \log n}{n}\right)$  data gathering capacity can be achieved as long as path loss exponent  $\alpha$  is such that  $2 < \alpha < 4$ ; indeed, for these values of  $\alpha$  the achieved capacity is shown to match the asymptotical upper bound, i.e., it is shown to be optimal. If  $\alpha > 4$ , the upper bound to capacity becomes  $O\left(\frac{W}{n}\right)$ , and optimal capacity can be achieved even without cooperative communication. Notice that although Zheng and Barton's results are based on less demanding assumptions than [3], they still build upon exploitation of advanced PHY layer techniques (cooperative communication), that might be difficult to use in a wireless sensor network.

The work that is more closely related to ours is [11], in which the authors investigate the data gathering capacity of a randomly deployed network in which  $s(n)$  out of the  $n$  nodes act as data sources,  $d(n)$  nodes are data sinks, and each source node is randomly mapped to a sink node. In this model, the authors identify two different asymptotic regimes: if  $d(n) = O\left(\sqrt{\frac{n}{\log n}}\right)$ , performance is constrained by the aggregate sink reception capacity, and per-source capacity is upper bounded by  $O\left(\frac{W \cdot d(n)}{s(n)}\right)$ ; if  $d(n) \gg \sqrt{\frac{n}{\log n}}$ , performance is constrained by network capacity, and per-source capacity saturates to  $O\left(\frac{W \cdot \sqrt{n/\log n}}{s(n)}\right)$  independently of the number  $d(n)$  of sink nodes. The authors also present a scheduling/routing algorithm to constructively lower bound capacity, and prove that the proposed construction indeed achieves asymptotically optimal data gathering capacity for most scaling regimes of  $s(n)$  and  $d(n)$ .

In this paper, we consider a traffic model similar to [11], in which  $s(n)$  nodes are data sources and  $d(n)$  nodes are data sinks. However, differently from [11], we assume that source/sink mapping obeys a locality criterion. In this model, we provide the following contributions:

- a) for the first time, we analyze not only data gathering capacity, but also *latency*. In particular, in some cases (see below) we present data gathering schemes simultaneously achieving asymptotical capacity *and* latency. We stress that turning a capacity achieving data gathering scheme into one that optimizes both capacity and latency is not a trivial task (see the construction used in Section IV).
- b) for the many-to-one case (i.e.,  $d(n) = 1$ ), we present a data gathering scheme that achieves asymptotically optimal data gathering capacity for arbitrary network deployments. To the best of our knowledge, ours is the first capacity-achieving construction for *arbitrary* network deployments. Furthermore, we show that, if a certain topological property that we call *minimal cell connectivity* is satisfied, the presented scheme achieves not only asymptotically optimal capacity, but also asymptotically optimal *latency* in data gathering.

totically optimal *latency* in data gathering.

- c) for the many-to-many case with arbitrary  $s(n)$  and  $d(n)$ , we consider both square grid and random network deployments, and present an extension of the data gathering scheme for single source networks that achieves asymptotically optimal data gathering latency. Furthermore, for most scaling regimes of  $s(n)$  and  $d(n)$ , the presented construction achieves also asymptotically optimal per-source data gathering capacity. In particular, the results for random network deployments show that per-source capacity can exceed the bounds established in [11] when  $\sqrt{\frac{n}{\log n}} \ll d(n) = O\left(\frac{n}{\log n}\right)$ . This is due to the different rule used to map data source to sinks: random in [11], locality-based in our study.
- c) finally, we consider the effect of a simple form of data aggregation based on a notion of sensing granularity, and show that per-source capacity can be increased by a factor  $f(n)$  with respect to the case of no data aggregation while preserving asymptotically optimal latency, where  $f(n)$  is node density.

Notably, and differently from most previous works, the presented results and data gathering scheme do not sacrifice energy efficiency to the need of maximizing capacity and minimizing latency. Furthermore, we want to stress that, differently from previous work on data gathering capacity, the results in this paper are derived using an accurate, SINR-based interference model, the physical interference model of [7].

### III. NETWORK MODEL AND PRELIMINARIES

We consider a wireless sensor network composed of  $n$  nodes, where  $s(n) \leq n - 1$  nodes act as data sources,  $d(n) \leq n - s(n)$  nodes act as data collection points (sinks), and the remaining  $n - s(n) - d(n)$  nodes may be used to relay data to the sinks. When multiple sink nodes are present in the network, the mapping between data sources and sinks obeys a locality criterion, which will be detailed in the following.

We assume nodes communicate through a common wireless channel of a certain, constant capacity  $W$ , and that the nodes transmission power is fixed to some constant value  $P$ . Correctness of packet reception is governed by a condition on SINR, according to the so-called physical interference model [7]. More specifically, when a node  $u$  transmits a packet to another node  $v$ , the packet is correctly received (at constant rate  $W$ ) if and only if

$$SINR(v) = \frac{P_v(u)}{N + \sum_{i \in \mathcal{T}} P_v(i)} \geq \beta,$$

where  $N$  is the background noise,  $\beta > 0$  is the capture threshold,  $\mathcal{T}$  is the set of nodes transmitting concurrently with node  $u$ , and  $P_v(x)$  is the received power at node  $v$  of the signal transmitted by node  $x$ . We make the standard assumption that radio signal propagation obeys the log-distance path loss model, according to which the received signal strength at distance  $d$  from the transmitter (for sufficiently large  $d$ , say,  $d \geq 1$ ) equals  $P \cdot d^{-\alpha}$ , where  $\alpha$  is the path loss exponent. In the following, we make the standard assumption that  $\alpha > 2$ , which is often the case in practice. Up to tedious technical details, our results can be generalized to the cost-based radio

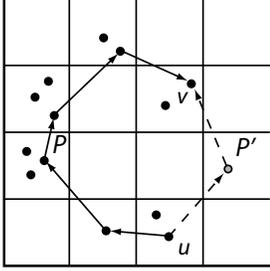


Fig. 1. The notion of minimal cell connectivity: without the gray node, the network is not minimally cell connected, since the length of minimal cell path  $P$  between source  $u$  and sink  $v$  is 5, and is greater than  $cd(u, v) = 2$ . With the gray node, the network becomes minimally cell connected, as  $len(P') = 2$ .

propagation model of [18], which is shown to closely resemble log-normal shadowing propagation.

For given values of  $P$ ,  $\beta$ ,  $\alpha$ , and  $N$ , we define the transmission range  $r_{max}$  of a node as the maximum distance up to which a receiver can successfully receive a packet *in absence of interference*. From the definition of physical interference model, we have  $r_{max} = \sqrt[3]{P/(\beta N)}$ .

The *communication graph*, denoted  $G = (\mathcal{V}, \mathcal{E})$ , represents the set of all possible communication links in the network, i.e.,  $\mathcal{V}$  is the set of the  $n$  nodes, and (undirected) edge  $(u, v) \in \mathcal{E}$  if and only if  $d(u, v) \leq r_{max}$ .

The *per-source data gathering capacity* of a network with  $n$  nodes,  $s(n)$  data sources, and  $d(n)$  sinks is defined as the maximum rate  $\lambda(n, s(n), d(n))$  at which each source node can send its data, with the property that the data sent by a source at time  $t$  is received by a sink node within time  $t + T$ , with  $T < \infty$ . The *data gathering latency* is the minimal time  $T(n, s(n), d(n))$  such that  $T \leq T(n, s(n), d(n))$ , for each data source. Note that, when  $d(n) = 1$ , the definitions of data gathering capacity and latency are meaningful only if graph  $G$  is connected. When  $d(n) > 1$ , it is indeed possible to have non-zero per-source capacity and finite latency also in partitioned networks. However, in order to be able to compare different data gathering scenarios on the same network topology, we will keep the assumption of connected communication graph also when  $d(n) > 1$ .

We consider two types of network deployment in this paper. In the *arbitrary* deployment, nodes are arbitrarily deployed in a two-dimensional region of arbitrary shape. In the *random* deployment, nodes are distributed uniformly at random in a square region of side  $(r_{max}/j) \cdot \sqrt{n/\log n}$ , where  $j > 8$  is an arbitrary constant. Note that in the random deployment scenario we can set the region side to be an arbitrary constant (e.g., 1), which implies that the transmission range  $r_{max}$  can be expressed as

$$r_{max} = r_{max}(n) = j \cdot \sqrt{\frac{\log n}{n}},$$

i.e., it is asymptotically vanishing with  $n$ . Thus, the random deployment scenario under the additional assumption that the side of the deployment region is a constant corresponds to the so called *dense random network* scenario, considered, for instance, in [11].

Similarly to [15], we assume the deployment region (in both arbitrary and random deployments) is partitioned into a lattice of square cells of side  $l$ , with  $l = \frac{r_{max}}{2h\sqrt{2}}$ , and  $h > 1$  an arbitrary constant. The step of the lattice ensures that, in absence of interference, a node located in a cell is able to communicate with *all* nodes in adjacent cells (horizontal, vertical, and diagonal adjacency). Note that condition  $h > 1$  ensures that the maximum distance between any two nodes in adjacent cells is at most  $\frac{r_{max}}{h}$ , i.e., strictly smaller than the transmission range. This is needed to guarantee that our capacity and latency achieving data gathering scheme, which only exploits transmissions between nodes in adjacent cells, avoid using long links with an arbitrarily small interference margin. These links, called black/gray links in [1], are notoriously difficult to simultaneously schedule in a STDMA scheme, and are thus avoided in our data gathering approach.

For a given node  $u$ ,  $C(u)$  denotes the cell to which node  $u$  belongs. Given any two nodes  $u, v$ , the *cell distance* between  $u$  and  $v$ , denoted  $cd(u, v)$ , is the minimum number of adjacent cells that must be traversed to reach  $C(v)$  starting from  $C(u)$  (or viceversa). Given a path  $\mathcal{P}$  in graph  $G$ , we say that  $\mathcal{P}$  is a *cell path* if and only if consecutive nodes in  $\mathcal{P}$  belongs to adjacent (non-empty) cells. Given any two nodes  $u, v$ , let  $\mathcal{P}(u, v)$  denote a minimal-length cell path between nodes  $u$  and  $v$ . We say that the network is *minimally cell connected* if and only if, for each source node  $u$ ,  $len(\mathcal{P}(u, v)) = cd(u, v)$ , where  $v$  is a sink node of minimal cell distance from node  $u$ <sup>1</sup>, and  $len(\mathcal{P})$  denotes the length of path  $\mathcal{P}$ . The notions of cell distance, cell path, and minimal cell connectivity are pictorially described in Figure 1. Observe that the notion of minimal cell connectivity, which we show to be sufficient to achieve asymptotically optimal data gathering latency, is closely related to the notion of backward connectivity, which is shown in [15] to be sufficient to achieve asymptotically optimal broadcast latency.

In the following, we assume that a simple form of data aggregation can be used to reduce the amount of data traffic circulating in the network. In our study, data aggregation is based on the notion of *sensing granularity*, which refers to the spatial resolution used for acquiring data from the sensed field. More specifically, we assume the network designer can specify a parameter  $sg$  defining sensing granularity, which is used to partition the network deployment region into non-overlapping square cells of side  $sg$ . When data aggregation is used, we assume that only *one* sensor data reading per sensing cell (if available) is to be returned to a sink node, independently of the actual number of nodes present in the sensing cell. Although very simple, this notion of data aggregation, which is based on the implicit assumption that the sensed field displays a high degree of spatial correlation, is of practical relevance. Note that the sensing cell partitioning is independent of the cell lattice used to define the notion of minimal cell connectivity and to design the data gathering scheme. To avoid confusion in terminology, in the following we use the term ‘‘cell’’ to refer to the cell used in the definition of the data gathering

<sup>1</sup>Minimal cell connectivity is trivially satisfied if both the source and the sink node reside in the same (or in adjacent) cell(s).

scheme, and the term ‘‘sensing cell’’ to refer to the cell used to aggregate data to be delivered to the sink(s).

We conclude this section with two known properties of the random network deployment, which will be useful in the remainder of this paper.

**Proposition 1.** (from [5]) *In the dense random deployment scenario, the critical transmission range for connectivity<sup>2</sup> is:*

$$ctr(n) = \frac{r_{max}(n)}{j}.$$

**Proposition 2.** (from [15]) *In the random deployment scenario, let  $C_{min}$  denote the number of nodes in the minimally occupied cell in the lattice. We have  $Prob(C_{min} > 0) \rightarrow 1$  as  $n \rightarrow \infty$ , i.e., w.h.p., each cell in the lattice contains at least one node.*

A straightforward consequence of Proposition 2 is that a randomly deployed network satisfies minimal cell connectivity, w.h.p.

#### IV. SINGLE SINK NETWORKS

##### A. No data aggregation

We first consider the case of single sink networks, i.e., throughout this section we assume  $d(n) = 1$ . Differently from most existing works on data gathering capacity, our results hold independently of the location of the sink node within the network.

We start by stating the upper bounds on data gathering capacity and latency, which hold for both arbitrary and random network deployments.

**Proposition 3.** *The per-source data gathering capacity of a single sink network with  $s(n)$  data sources is  $O\left(\frac{W}{s(n)}\right)$ , where  $W$  is the capacity of the wireless link. The data gathering latency is  $\Omega\left(\frac{D(n)}{r_{max}}\right)$ , where  $D(n)$  is the network diameter.*

The upper bound on capacity follows by observing, similarly to [11], [12], that the sink node can receive data from at most one node at any instant of time, that the channel data rate is  $W$ , and that sink reception capacity must be shared amongst  $s(n)$  sources. We thus do not consider use of advanced PHY layer techniques to exceed the  $O(W)$  receiving capacity limit at the sink node, as proposed in [4], [24]. The lower bound on latency is based on the observation that, under our network model, a packet can travel at most distance  $r_{max}$  towards the sink node at each communication round.

We now present the capacity and latency achieving data gathering scheme. We assume that each node  $v$  is aware of the cell  $C(v)$  to which it belongs, and that a spatial TDMA approach is used at the MAC layer. Furthermore, we assume that cells are colored according to a two-dimensional coloring scheme (see Figure 2), similar to those used, e.g., in [8], [15]. More specifically, the color of a cell (and of all the nodes belonging to that cell) is a tuple  $(i, j)$ , with  $i$  the horizontal color and  $j$  the vertical color. A total of  $k^2$  colors,

<sup>2</sup>The critical transmission range for connectivity is defined as the minimum value of the transmission range such that the resulting communication graph is connected with high probability, i.e., with probability approaching 1 as  $n$  grows to infinity.

(2,0)	(0,0)	(1,0)	(2,0)	(0,0)	(1,0)	(2,0)	(0,0)	(1,0)	(2,0)	(0,0)	(1,0)
(2,2)	(0,2)	(1,2)	(2,2)	(0,2)	(1,2)	(2,2)	(0,2)	(1,2)	(2,2)	(0,2)	(1,2)
(2,1)	(0,1)	(1,1)	(2,1)	(0,1)	(1,1)	(2,1)	(0,1)	(1,1)	(2,1)	(0,1)	(1,1)
(2,0)	(0,0)	(1,0)	(2,0)	(0,0)	(1,0)	(2,0)	(0,0)	(1,0)	(2,0)	(0,0)	(1,0)
(2,2)	(0,2)	(1,2)	(2,2)	(0,2)	(1,2)	(2,2)	(0,2)	(1,2)	(2,2)	(0,2)	(1,2)
(2,1)	(0,1)	(1,1)	(2,1)	(0,1)	(1,1)	(2,1)	(0,1)	(1,1)	(2,1)	(0,1)	(1,1)
(2,0)	(0,0)	(1,0)	(2,0)	(0,0)	(1,0)	(2,0)	(0,0)	(1,0)	(2,0)	(0,0)	(1,0)

Fig. 2. Two-dimensional coloring scheme used in the data gathering scheme with  $k = 3$ . Cells with the same color  $(0, 0)$  are shaded.

for some constant  $k > 0$ , are used, i.e., both the horizontal and vertical colors take values in  $\{0, \dots, k-1\}$ . The coloring scheme is used in an obvious way to form the transmission schedule, i.e., time is subdivided into slots of equal time duration, which are numbered in a cyclical fashion between 0 and  $k^2 - 1$ ; each color is assigned a transmission opportunity at each *communication round* composed of  $k^2$  consecutive transmission slots. When the slot corresponding to color  $(i, j)$  is active, only links whose transmitter is in a cell of color  $(i, j)$  can be scheduled for transmission.

The following lemma, proved in [15], is a fundamental building block of our construction:

**Lemma 1.** *Assume a cell partitioning as defined in Section III, and assume that at most one node per cell with the same color transmits. If*

$$k \geq \bar{k} = \left\lceil 2 + 2^{\frac{3}{2} + \frac{4}{\alpha}} (\beta \zeta(\alpha - 1) h^\alpha / (h^\alpha - 1))^{\frac{1}{\alpha}} \right\rceil,$$

where  $\zeta$  is the Riemann’s zeta function, then the packet sent by a node is correctly received by all nodes in adjacent cells.

The lemma states that a *constant* degree of spatial separation (number of colors) is sufficient to achieve successful packet reception in neighboring cells, under the assumption that at most one node per cell with the same color is simultaneously transmitting. Thus, a constant degree of spatial separation is sufficient to achieve  $\Omega(r_{max})$  (recall that the size of the cell is  $\Theta(r_{max})$ ) progress of a packet towards the sink (if minimal cell connectivity is satisfied).

The data gathering scheme is reported in Figure 3. Depending on its role, each network node is classified as either sink, source node, or relay node. The sink node is the only data collector in the network. Source nodes generate new data packets to be delivered to the sink; furthermore, they can act as data forwarders if they are not sending their own data. Relay nodes can act only as forwarders of data generated by source nodes. Non-sink nodes (either source or relay nodes) can be selected as *leader* nodes in their cell. There is a unique leader node in each non-empty cell, which is selected according to some rule, such as at random, remaining battery level, etc. Leader nodes act as forwarders of data towards the sink. Note that only relay nodes that are cell leaders participate in forwarding the collected data to the sink. Non-leader relay nodes actually do not take part in the data gathering scheme, and can be considered as non-active nodes (and possibly put into sleep mode to save energy).

The data gathering scheme is very simple. Non-sink nodes are allowed to transmit only if the current slot color corre-

Algorithm for a generic non-sink node  $v$ :

- Let  $i$  be the color of the current time slot  
 Node  $v$  has a single position packet buffer
1. **if**  $color(v) = i$  **then**
  2.   **if**  $source(v)$  **and**  $MyTurn(v)$  **then**  
     generate and transmit new packet
  3.   **else if**  $cellLeader(v)$  **then**
  4.     **if**  $buffer(v)$  is not empty **then**
  5.       transmit packet and empty buffer
  6. **else** //  $color(v) \neq i$
  7. **if**  $cellLeader(v)$  **then**
  8.   listen to the channel
  9.   **if** new packet arrive **then**
  10.   receive the packet and store in transmit buffer

Algorithm for the sink node:

- 1s. At each time slot:
- 2s.   listen to the channel
- 3s.   **if** new packet arrive **then**
- 4s.     receive the packet and deliver to user application

Fig. 3. The data gathering scheme with single sink node.

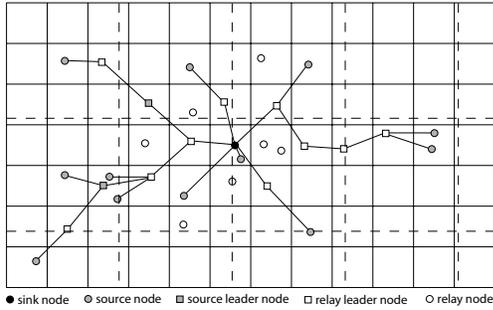


Fig. 4. The data gathering tree. Sensing cells are shown with dashed lines. If data aggregation is used, only one source node per sensing cell is active.

sponds to the node color. If the node is a source, a further condition must be satisfied in order to enable transmission of a new packet. This condition, which is verified by function  $MyTurn()$  at step 2., ensures that data generated by the various sources do not conflict at intermediate nodes in their route to the sink. In the following, we show the constructive technique used to calculate function  $MyTurn()$  for each source node, and prove its correctness. If transmission of a new data packet is not allowed, the node can act as forwarder of data generated by other sources if it is a leader node (step 3.). If the current slot color does not correspond to the node color, the node actively participates in the data gathering scheme only if it is a leader. In this case, it checks whether a new packet is received at the current time slot, and stores the new packet in the buffer. The data gathering algorithm for the sink node is straightforward.

To ensure collision-free delivery of data to the sink and optimal latency (in case minimal cell connectivity is satisfied), data is conveyed to the sink through a cell-based shortest path tree, called the *data gathering tree*, where the transmission time (slot) for each data source is appropriately computed<sup>3</sup>. An example of data gathering tree is reported in Figure 4: for each source node  $v$ , a minimal length cell path to sink node

<sup>3</sup>A similar construction and transmission time computation has been used in the broadcast scheme for arbitrary number of sources presented in [16].

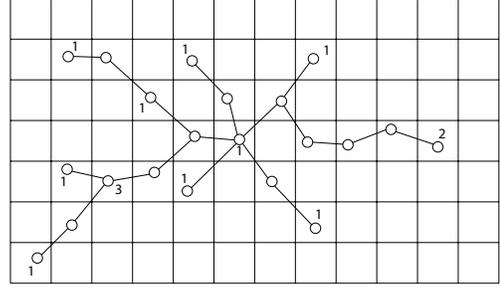


Fig. 5. The cell tree corresponding to the data gathering tree of Figure 4. Only positive cell-node weights are shown.

$s$ , denoted  $\mathcal{P}(v, s)$ , is randomly selected as data forwarding path<sup>4</sup>. Note that, without loss of generality, we can assume that  $\mathcal{P}(v, s) - \{v, s\}$  is composed only of leader nodes. The composition of all these minimal length cell paths (one for each source) forms the data gathering tree.

Starting from the data gathering tree, a simplified gathering tree, called *cell tree*, is constructed as follows. A cell-node is associated with each cell with at least one node in the data gathering tree. With a slight abuse of notation, in the following we use  $C(T)$  to denote the cell corresponding to cell-node  $T$  in the cell tree. An edge is present between cell-nodes  $T$  and  $U$  if and only if there is at least one edge between a node in  $C(T)$  and a node in  $C(U)$  in the data gathering tree. Each cell-node is assigned with a weight equal to the number of source nodes present in the corresponding cell. The cell tree corresponding to the data gathering tree of Figure 4 is reported in Figure 5.

The main idea of the data gathering scheme is to have the sink node receiving a new packet from a different data source at each communication round. Since the communication round is composed of a constant number  $\bar{k}^2$  of transmission slots, and the total number of sources is  $s(n)$ , we have that each source delivers packets to the sink with a rate of  $\frac{W}{\bar{k}^2 \cdot s(n)} = \Omega(\frac{W}{s(n)})$ , which is asymptotically optimal. Furthermore, data source transmission times are set in such a way that conflict at intermediate nodes of the cell tree are avoided, which ensures that a packet transmitted by source  $v$  during communication round  $t$  is received by sink  $s$  at communication round  $t + cd(v, s)$ . As we shall see, this property implies asymptotically optimal data gathering latency when minimal cell connectivity of the network is satisfied.

The transmission time for each source node in the network is recursively computed as follows, starting from the root of the cell tree, which corresponds to the cell containing the sink node. Communications rounds (each composed of  $\bar{k}^2$  transmission slots) are cyclically numbered from 0 to  $s(n) - 1$ . In the following, all the mathematics is modulo  $s(n)$ .

Let  $T_1, \dots, T_h$  be the children of the root node  $R$  in the cell tree, and let  $sw(T_i)$  be the sum of the weights of all cell-nodes belonging to the subtree rooted at  $T_i$ . Furthermore, denote by

<sup>4</sup>The choice of the actual minimal length cell path used to deliver data to the sink, which determines the tree topology, has no influence on the data gathering capacity and latency achieved by our scheme.

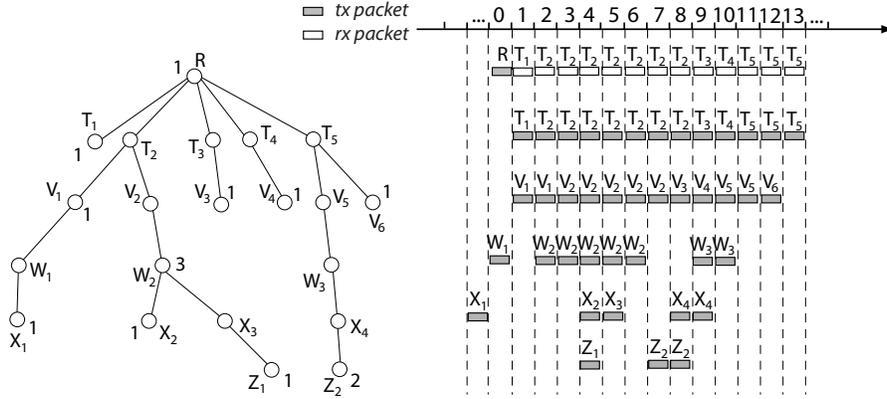


Fig. 6. Transmission scheduling for the cell tree of Figure 5.

$w(T)$  the weight of cell-node  $T$ . It is easy to see that

$$w(R) + \sum_{i=1}^h sw(T_i) = s(n).$$

Without loss of generality, we let the communication rounds  $\{0, \dots, w(R)-1\}$  be assigned to the  $w(R)$  sources in the same cell as the sink (if any), and communication rounds

$$\left\{ \begin{aligned} L_i &= w(R) + \sum_{j=1}^{i-1} sw(T_j), \dots \\ \dots, w(R) + \sum_{j=1}^i sw(T_j) - 1 &= U_i \end{aligned} \right\}$$

be assigned to the sources that gather their data to the sink via cell-node  $T_i$ , with  $i = 1, \dots, h$ . We can now compute function  $MyTurn()$  for source nodes in  $C(s)$  by randomly ordering them, and letting  $MyTurn(v)$  return true for a source node  $v$  in  $C(s)$  and order position  $i$  if and only if the current communication round is  $i - 1$ .

Let us now consider one of the children nodes  $T_i$ . If  $w(T_i) > 0$ , i.e., if at least one data source is located in  $C(T_i)$ , then the first  $w(T_i)$  rounds of the  $[L_i, U_i]$  interval are assigned to sources in  $C(T_i)$ , and function  $MyTurn()$  for these sources can be computed as above. The remaining rounds are recursively assigned to the children  $T_1^i, \dots, T_{l_i}^i$  of cell-node  $T_i$  in the tree, using an interval-based round assignment as above. However, we have to consider that node  $T_i$  must be able to *transmit* a new packet during each of its assigned rounds. Hence, the packets coming from a particular children  $T_j^i$ , which must be transmitted by the leader node of cell  $C(T_i)$  during interval  $[L_j^i, U_j^i]$ , must be received by  $T_i$  in an earlier time window. In order to preserve optimality of data gathering latency and single buffering at the nodes, we set the time window used by cell-node  $T_j^i$  to transmit its data to cell-node  $T_i$  to be  $[L_j^i - 1, U_j^i - 1]$ . In other words, the data transmitted by the leader node in cell  $C(T_j^i)$  at round  $t$  is received by the leader node in cell  $C(T_i)$  during the same round, and transmitted to the parent cell-node during round  $t + 1$ .

By repeatedly applying the above construction, the transmission times of all data sources in the network can be computed, and used to define the respective  $myTurn()$  functions. The result of the transmission time computation for the data

gathering tree reported in Figure 4 is reported in Figure 6. Note that it is perfectly possible that, as a consequence of schedule computation, several nodes simultaneously transmit packets within the cell tree. However, the data gathering scheme ensures that at most one node per cell is allowed to transmit in each transmission slot which, by Lemma 1, implies that each simultaneously transmitted packet can be successfully received by nodes in neighboring cells as long as  $k \geq \bar{k}$ .

We are now ready to prove that our data gathering scheme provides asymptotically optimal capacity.

**Theorem 1.** *The data gathering scheme of Figure 3 achieves  $\Omega\left(\frac{W}{s(n)}\right)$  per-source capacity.*

*Proof:* The proof follows easily by observing that: *i)* the transmission schedule is built in such a way that packets generated by any source are delivered to the sink once every  $s(n)$  communications rounds; *ii)* the link communication rate is  $W$ ; and *iii)* each communication round is composed of a constant number  $\bar{k}^2$  of fixed length transmission slots, where the value of  $\bar{k}$  is defined in Lemma 1. ■

The following Lemma provides an upper bound to data gathering latency as a function of the cell distance to the sink, under the assumption that the network is minimally cell connected.

**Lemma 2.** *Assume the network is minimally cell connected. The packet sent by source node  $v$  during communication round  $t$  is received by the sink  $s$  during communication round  $t + cd(v, s)$ .*

*Proof:* The transmission schedule is built in such a way that every generated data packets traverses one edge of the cell tree at each communication round. Hence, the packet generated at round  $t$  by source node  $v$  such that the corresponding cell-node is at distance  $h$  from the root of the cell tree is received by the sink node at round  $t + h$ . The cell tree is built starting from the data gathering tree, which is in turn constructed by joining minimal cell paths for each data source. Given minimal cell connectivity, we have that, for each source, the minimal cell path to the sink has length equal to  $cd(v, s)$ . Hence, it must be  $h = cd(v, s)$ , and the lemma is proved. ■

**Theorem 2.** *Assume the network is minimally cell connected.*

The data gathering scheme of Figure 3 achieves  $O\left(\frac{D(n)}{r_{max}}\right)$  data gathering latency.

*Proof:* We observe that, since  $\bar{k}$  is a constant and the duration of a transmission slot is fixed, the duration of a communication round does not depend on  $n$ . This implies also that the actual transmission slot during which the packet is transmitted by the source (or received by the sink node), which can add at most an additional factor equal to the duration of a communication round to latency, is  $O(1)$ . By Lemma 2, and by the above observation, the data gathering latency for source  $v$  is  $O(cd(v, s))$ . We next observe that, by definition of network diameter, the Euclidean distance between any source node  $v$  and the sink  $s$  is upper bounded by  $D(n)$ . The proof of the theorem follows by observing that the cell size is  $\Theta(r_{max})$ , which implies that  $cd(v, s) = O\left(\frac{D(n)}{r_{max}}\right)$ . ■

Combining Proposition 3 and theorems 1 and 2, we can conclude the following:

**Theorem 3.** *The data gathering scheme of Figure 3 achieves asymptotically optimal data gathering capacity and latency. This result holds for arbitrary network deployments, subject to the condition that minimal cell connectivity is satisfied.*

### B. Data aggregation

We now consider the effects of data aggregation on data gathering performance. To this purpose, we assume the deployment region is divided into square sensing cells of side  $sg$ . The data gathering scheme is modified as follows. If more than one data sources are present in a sensing cell, only one of them is selected as *active source* (an arbitrary criterion can be used to select active sources); the data gathering scheme remains unchanged, with the difference that only active sources are allowed to transmit and used to build the data gathering tree.

It is immediate to see that data gathering latency is not impacted by data aggregation. To evaluate the effects of data aggregation on data gathering capacity, we need to determine the relationship between  $sg$  and the area of the deployment region. It is in fact immediate to see that per-source data gathering capacity in presence of data aggregation becomes

$$\Omega\left(\frac{W}{as(n)}\right),$$

where  $as(n)$  is the number of active sources in the network. In turn,  $as(n)$  is upper bounded by the number of sensing cells in the network, which can be determined only if the relationship between  $sg$  and the area of the deployment region is fixed. To this purpose, we consider two deployment scenarios, similar to those considered in previous works on the data gathering capacity [4], [11], [12]. In the *square grid* scenario, the  $n$  nodes are distributed in a square lattice of step  $\Theta(r_{max})$ ; thus, the deployment region is a square of side  $L = \Theta(\sqrt{n} \cdot r_{max}) = \Theta(\sqrt{n})$ . Defining the *density* of a deployment as the average number of nodes within a node transmission range, it is easy to see that square grid deployment is a minimal density scenario that maintains the network connected. The other considered scenario is the random deployment scenario described in Section III. We recall that, under this scenario,  $n$  nodes are

distributed uniformly at random in a square region of side  $L = (r_{max}/j) \cdot \sqrt{n/\log n}$ . As shown in Section III, the random deployment scenario is a minimal density scenario for achieving connectivity w.h.p. with random uniform node deployment. Note that, due to random node distribution, this minimal density is  $\Theta(\log n)$ , which is larger than the  $\Theta(1)$  density of the square grid scenario.

Let us first consider the square grid deployment; in this case, the number of sensing cells is

$$\Theta\left(\left(\frac{\sqrt{n}}{sg}\right)^2\right) = \Theta(n),$$

under the reasonable assumption that the sensing granularity does not depend on the number of nodes. Thus, in the square grid deployment, owing to the low node density, data aggregation provides no significant gain in per-source data gathering capacity, which remains  $\Omega\left(\frac{W}{n}\right)$ .

Let us now consider the random deployment; in this case, the number of sensing cells is

$$\Theta\left(\left(\sqrt{\frac{n}{\log n}} \cdot \frac{1}{sg}\right)^2\right) = \Theta\left(\frac{n}{\log n}\right),$$

which implies that the per-source data gathering capacity is increased of a  $\Theta(\log n)$  factor over the case of no data aggregation. It is easy to generalize the above reasoning, and conclude that, under the assumption that the density of node deployment is  $\Theta(f(n))$ , for some  $f(n) = O(n)$ , the per-source data gathering capacity is increased by a  $f(n)$  factor.

We can then conclude this section with the following theorem:

**Theorem 4.** *Assume a network is deployed with  $\Theta(f(n))$  density, for some  $f(n) = O(n)$ ; furthermore, assume that the network is minimally cell connected, and that the sensing granularity does not depend on  $n$ . Then, our data gathering scheme combined with data aggregation achieves optimal data gathering latency, and  $\Omega\left(\frac{W \cdot f(n)}{n}\right)$  per-source data gathering capacity.*

## V. MULTIPLE SINK NETWORKS

We now consider the more general case of multiple sink networks, where  $d(n) > 1$  nodes in the network act as data collector. When multiple sinks are present in the network, a relevant question is how to map data source to sinks. In [11], the authors assume this mapping is random. This assumption, which eases the theoretical analysis developed in [11], clearly leads to sub-optimal data gathering performance, since the number of re-transmissions (which negatively impact per-source data gathering capacity) needed to deliver a data packet increases with the distance to the sink. Driven by this observation, in this section we make the assumption that source nodes send their data to a sink at short cell distance. Furthermore, we assume that sink nodes are not randomly located (as in the model of [11]), but they can be properly selected amongst the network nodes, as it is typically the case in a wireless sensor network scenario. On the other hand, source nodes are assumed to be randomly chosen amongst the remaining  $n - d(n)$  nodes.

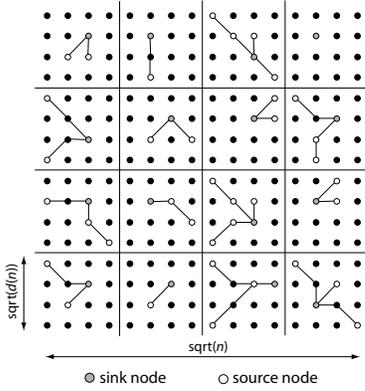


Fig. 7. Data gathering scheme with multiple sinks in square grid deployments.

We first state the capacity upper bound and latency lower bound in presence of  $s(n)$  source nodes and  $d(n)$  sinks.

**Proposition 4.** *The per-source data gathering capacity of a network with  $d(n)$  sinks and  $s(n)$  sources is  $O\left(\frac{W \cdot d(n)}{s(n)}\right)$ . The data gathering latency is  $\Omega\left(\frac{D(n)}{r_{max} \cdot \sqrt{d(n)}}\right)$ .*

The capacity bound follows by observing that, with  $d(n)$  sink nodes, the aggregate sink reception capacity is increased by a factor at most  $d(n)$  with respect to the single source case. The lower bound on latency is obtained under the optimistic assumption that sink nodes are evenly spread in the network, so that the distance of a source to the closest sink can be reduced by a factor  $O(\sqrt{d(n)})$ .

In the remainder of this section, we consider two possible network deployments, namely the square grid deployment described at the end of Section IV and the random deployment.

Assume square grid node deployment. The deployment region is divided into  $d(n)$  square sub-regions of approximately equal size<sup>5</sup>, and one node is selected as sink node within each sub-region (note that at least one node is contained in each sub-region). Source nodes belonging to a certain sub-region  $R$  send their data towards the only sink node within  $R$  (see Figure 7). The data gathering scheme described in the previous section is independently applied in each sub-region, i.e., a data gathering and cell tree is built in each sub-region, and the data gathering process goes on in parallel along all the data gathering trees. We now characterize the data gathering capacity and latency achieved by this data gathering approach:

**Theorem 5.** *Assume square grid node deployment. The multiple sink data gathering scheme described above achieves  $\Omega\left(\frac{W}{M(s(n), d(n))}\right)$  per-source data gathering capacity, where  $M(s(n), d(n))$  denotes the maximum number of source nodes in a sub-region, and  $O\left(\frac{D(n)}{r_{max} \cdot \sqrt{d(n)}}\right)$  (i.e., asymptotically optimal) data gathering latency.*

*Proof:* We first observe that, even if data gathering on the different trees goes on in parallel, transmissions in the different trees do not corrupt each other. In fact, since each cell

<sup>5</sup>To simplify presentation, in the following we assume that each cell as defined in Section III is entirely contained in a single sub-region. This assumption might entail some rounding on the actual size of sub-regions.

is entirely contained in a single sub-region, the property that at most one node per cell with the same color is simultaneously transmitting holds also when multiple, spatially separated data gathering processes are active. Thus, Lemma 1 is satisfied, and source nodes in sub-region  $R$  can transmit their data to the respective sink with rate  $\frac{W}{s(R)}$ , where  $s(R)$  is the number of sources in sub-region  $R$ . It is then clear that per-source data gathering capacity is lower bounded by  $\frac{W}{M(s(n), d(n))}$ , where  $M(s(n), d(n)) \geq S(R)$  is the maximum number of sources in a sub-region. The upper bound on data gathering latency follows immediately by: *i*) construction of sub-regions; *ii*) construction of the data gathering trees in each sub-region; and *iii*) the observation that the square grid deployment satisfies minimal cell connectivity. ■

The value of  $M(s(n), d(n))$ , and thus the achievable per-source capacity, depends on the relative magnitude of  $s(n)$  and  $d(n)$ , and is characterized in Lemma 1 of [11]. In particular, if  $s(n) \gg d(n) \log(d(n))$  (e.g.,  $s(n) = n^a$  and  $d(n) = n^b$ , with  $0 < b < a < 1$ ), we have  $M(s(n), d(n)) = \Theta\left(\frac{s(n)}{d(n)}\right)$ , which implies that our scheme achieves asymptotically optimal  $\Omega\left(\frac{W \cdot d(n)}{s(n)}\right)$  per-node capacity. Another relevant scenario is when  $s(n) = d(n) = \Theta(n)$ , i.e., when the order of the number of sources is the highest possible, and a comparable number of sinks is used. In this case, we have  $M(s(n), d(n)) = \Theta\left(\frac{\log n}{\log \log n}\right)$ , and the achieved per-node capacity is  $\Omega\left(\frac{W \log \log n}{\log n}\right)$ , which is a factor  $\frac{\log n}{\log \log n}$  below optimal. The sub-optimal behavior of our data gathering scheme is due to unbalanced distribution of source nodes in the sub-regions.

Consider now the case of random node deployment. Similarly to the previous case, we subdivide the deployment region into  $d(n)$  sub-regions of approximately equal side. However, we add the further constraint that sub-regions must be larger than a cell, as defined in Section III. This is to ensure that, in accordance with Proposition 2, each sub-region contains at least one node w.h.p., so that a sink node can be selected within each sub-region w.h.p. Note that this constraint imposes an upper bound on the number of admissible sink nodes, namely  $d(n) = O\left(\frac{n}{\log n}\right)$ .

The data gathering scheme is the same as in the case of square grid deployment: one node is selected as sink in each of the  $d(n)$  sub-regions; all source nodes in sub-region  $R$  send their data to the sink node belonging to  $R$ ; data gathering along the  $d(n)$  data gathering trees runs in parallel. The following theorem can be proved along the same lines as Theorem 5:

**Theorem 6.** *Assume random node deployment, and that  $d(n) = O\left(\frac{n}{\log n}\right)$ . The multiple sink data gathering scheme described above achieves  $\Omega\left(\frac{W}{M(s(n), d(n))}\right)$  per-source data gathering capacity, where  $M(s(n), d(n))$  denotes the maximum number of source nodes in a sub-region, and  $O\left(\frac{D(n)}{r_{max} \cdot \sqrt{d(n)}}\right)$  (i.e., asymptotically optimal) data gathering latency.*

We observe that when  $d(n) = \Theta\left(\frac{n}{\log n}\right)$  and  $s(n) = \Theta(n)$ ,

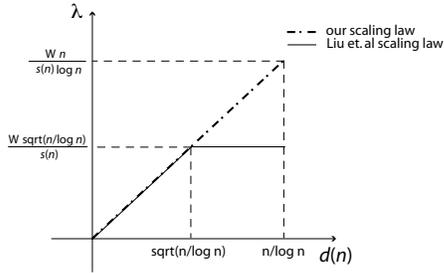


Fig. 8. Data gathering capacity scaling laws with random network deployment.

we have  $M(s(n), d(n)) = \frac{s(n)}{d(n)}$ , which implies that the per-source capacity  $\Omega\left(\frac{W \cdot d(n)}{s(n)}\right) = \Omega\left(\frac{W}{\log n}\right)$  achieved by our scheme is optimal. It is interesting to compare this result with the findings of [11], which also assume random network deployment. In [11], the authors show that per-source capacity is upper bounded by  $O\left(\frac{W \cdot d(n)}{s(n)}\right)$  when  $d(n) = O\left(\sqrt{\frac{n}{\log n}}\right)$ , while it is upper bounded by  $O\left(\frac{W \cdot \sqrt{n/\log n}}{s(n)}\right)$  for larger values of  $d(n)$ . Conversely, our results have shown that  $O\left(\frac{W \cdot d(n)}{s(n)}\right)$  per-source capacity can be achieved even when  $d(n) \gg \sqrt{\frac{n}{\log n}}$ , as long as  $d(n) = O\left(\frac{n}{\log n}\right)$ . This different scaling law in per-source capacity, which is depicted in Figure 8, is due to the different strategies used to map source node to sinks: random in [11], locality-based in our work. Thus, the advantages of the locality criterion used for data gathering becomes tangible in terms of achievable capacity.

Before ending this section, we observe that data aggregation can be used also in presence of multiple sink to increase per-source capacity by a factor of  $f(n)$ , where  $f(n)$  is the node density.

## VI. ENERGY EFFICIENCY

Differently from most previous work on data gathering capacity characterization, the capacity/latency achieving data gathering scheme presented in this paper is fully compatible with an energy efficient wireless sensor network design. In particular, energy-efficiency is achieved for what concerns choice of the *i*) node transmission range, *ii*) data aggregation method, and *iii*) data source/sink mapping.

Concerning *i*), we observe that Proposition 1 states that, up to a constant factor, the transmission range as defined in the dense random deployment scenario considered in this paper is the *minimum possible to maintain the network connected*, w.h.p. In other words, choosing a smaller value for the transmission range would result in a disconnected network with some fixed probability  $p > 0$ , impairing a fundamental network property. Thus, the results derived in this paper are obtained considering the minimum possible (i.e., energy efficient) setting for the transmission range. We stress that this is not the case in some relevant related work on data gathering. For instance, in [4] the cooperative transmission scheme requires using very high transmit power values at the neighboring nodes to achieve an  $\Omega(W \log n)$  capacity at the sink node. In [12], the lower bound on the per-source data

gathering capacity is of the form

$$\lambda(n, n-1, 1) \geq \frac{W}{n} \cdot \frac{c_1 r_{max}^2 - c_2}{c_3 r_{max}^2 + c_4 r_{max} + c_5}, \quad (1)$$

with the  $c_i$ s positive constants. The above lower bound is  $\Omega\left(\frac{W}{n}\right)$ , i.e., asymptotically optimal, if  $r_{max}$  does not depend on  $n$ . However, in [12] the authors consider a dense network scenario in which an increasingly large number of nodes is deployed in a fixed area. Hence, an energy efficient setting of the transmission range would require  $r_{max} \rightarrow 0$  as  $n \rightarrow \infty$  (as in our analysis), to take advantage of the increasing node density. However, the capacity lower bound reduces to the trivial lower bound  $\lambda(n, n-1, 1) \geq 0$  when  $r_{max} \rightarrow 0$ , since the right hand side of (1) converges to a negative constant when  $r_{max} \rightarrow 0$ . Thus, the optimal capacity lower bound must be traded off with energy efficiency in [12], which is not the case for the results presented in this paper.

Concerning *ii*), we observe that, although very simple, the notion of data aggregation used in our analysis is of practical relevance, and fits well with sleep scheduling techniques used to extend sensor network lifetime without significantly impacting observation accuracy (see, e.g., [2], [22]).

Finally, concerning *iii*) we observe that the locality rule used to map active data sources to sinks is fully compatible with energy-efficient routing algorithms, in which a minimal number of hops (and, hence, packet retransmissions) is needed for an active source to reach a sink. This is in sharp contrast with the random mapping used in [11], which on the average leads to energy inefficient routing of packets to the sinks.

## VII. DISCUSSION AND FINAL CONSIDERATIONS

In this paper, we have investigated the capacity and latency scaling laws of data gathering in wireless sensor networks where  $s(n)$  sources have to send their data to one of the  $d(n)$  sinks. Our main finding is that a simple gathering scheme can be used to obtain optimal or near optimal data gathering capacity for most scaling regimes of  $s(n)$  and  $d(n)$ . Furthermore, the same gathering scheme achieves optimal data gathering latency under the condition that the network is minimally cell connected, a condition which is satisfied, for instance, in case of square grid and random network deployment. We have also characterized the capacity increase achieved by a form of data aggregation based on the notion of sensing granularity, increase which is shown to be proportional to the node density  $f(n)$ . Although simple, this form of data aggregation is of practical relevance, and fits well with sleep scheduling strategies aimed at extending network lifetime.

Before ending this paper, we comment on the following two questions: *i*) what are the design guidelines that can be derived from our theoretical results? and *ii*) is the characterized data gathering capacity sufficient to guarantee observability of the region monitored by a wireless sensor network?

Concerning *i*), we observe that the capacity and latency achieving data gathering scheme presented in this paper is fully distributed and very easy to implement in a real network. However, it is based on a set of assumptions, which might not always hold in a wireless sensor network. More specifically, it is assumed that spatial TDMA is used at the MAC layer (which

requires time synchronization), a weak form of location-awareness (nodes must know the cell to which they belong), and that leader nodes are selected in each cell. Indeed, time synchronization (see, e.g., [21]), location-awareness (see, e.g., [14]), and leader election mechanisms (see, e.g., [3]) are quite typical building blocks of a wireless sensor network used, e.g., for environmental monitoring, so we believe that the proposed data gathering scheme can potentially be implemented in several wireless sensor network application scenarios.

As for *ii*), we observe that the optimal per-source capacity without data aggregation is at most  $O\left(\frac{W \cdot d(n)}{s(n)}\right)$ , i.e., asymptotically vanishing unless  $d(n) = \Omega(s(n))$ . Thus, the only way of achieving non-vanishing per-source capacity is to have a number of sinks comparable to the number of sources, which might entail a prohibitive cost in many application scenarios. However, if the observed field is spatially correlated, data aggregation can be used to increase per-source capacity by selecting a number of available sources to report their data, and having a large set of nearby sources to remain silent. The advantage of data aggregation over the case of no aggregation clearly depends on the (source) node density: the higher the density, the higher capacity increase can be achieved. It is interesting to observe that in random deployments, the minimal density necessary to guarantee network connectivity w.h.p., which is  $\Omega(\log n)$ , achieves the same  $O(\log n)$  capacity increase (in the single sink scenario) that has been achieved in [4], [24] through sophisticated PHY layer and data encoding techniques. The resulting per-source capacity has been shown in [4] to be sufficient to guarantee observability of a random Gaussian field. Although this does not directly imply that the same observability property holds also in our setting (in fact, nodes are evenly spaced on a sphere in the model of [4], while they are randomly deployed in a square in our case), the above observation seems to indicate that a simple form of data aggregation has the potential to guarantee observability of a random field also in presence of a single sink.

It is also worth observing that our results and the presented data gathering scheme are fully compatible with an energy efficient network design in which, for instance, the transmission range is set to the minimum possible value that keeps the network connected, a subset of the nodes are put in sleep mode to save energy, and so on.

Finally, we want to outline that the results presented in this paper are relevant also to other types of wireless multihop networks where the many-to-one or many-to-many traffic pattern is likely to arise, such as wireless mesh networks where most of the traffic is directed towards few gateway nodes.

#### REFERENCES

- [1] D. Blough, G. Resta, P. Santi, "Approximation Algorithms for Wireless Link Scheduling with SINR-based Interference", *Tech. Rep. IIT-TR03/2008*, Istituto di Informatica e Telematica, Pisa, Italy.
- [2] Q. Cao, T. Abelzahr, T. He, J. Stankovic, "Towards Optimal Sleep Scheduling in Sensor Networks for Rare-Event Detection", *Proc. Int. Symposium on Information Processing in Sensor Networks (IPSN)*, pp. 20–27, 2005.
- [3] Q. Dong, D. Liu, "Resilient Cluster Leader Election for Wireless Sensor Network", *Proc. IEEE Secon*, pp. 108–116, 2009.
- [4] H. El Gamal, "On the Scaling Laws of Dense Wireless Sensor Networks: The Data Gathering Channel", *IEEE Trans. on Information Theory*, Vol. 51, n. 3, pp. 1229–1234, 2005.

- [5] R.B. Ellis, J.L. Martin, C. Yan, "Random Geometric Graph Diameter in the Unit Ball", *Algorithmica*, Vol. 47, n. 4, pp. 421–438, 2007.
- [6] M. Grossglauser, D.N.C. Tse, "Mobility Increases the Capacity of Ad-Hoc Wireless Networks", *Proc. IEEE Infocom*, pp. 1360–1369, 2001.
- [7] P. Gupta and P.R. Kumar, "The Capacity of Wireless Networks," *IEEE Transactions on Information Theory*, Vol. 46, No. 2, pp. 388–404, 2000.
- [8] A. Keshavarz-Haddad, R. Riedi, "On the Broadcast Capacity of Multihop Wireless Networks: Interplay of Power, Density and Interference", *Proc. IEEE SECON*, pp. 314–323, 2007.
- [9] A. Keshavarz-Haddad, V. Ribeiro, R. Riedi, "Broadcast Capacity in Multihop Wireless Networks", *Proc. ACM Mobicom*, pp. 239–250, 2006.
- [10] X.Y. Li, S.J. Tang, O. Frieder, "Multicast Capacity for Large Scale Wireless Ad Hoc Networks", *Proc. ACM Mobicom*, pp. 266–277, 2007.
- [11] B. Liu, D. Towsley, A. Swami, "Data Gathering Capacity of Large Scale Multihop Wireless Networks", *Proc. IEEE MASS*, pp. 124–132, 2008.
- [12] D. Marco, E.J. Duarte-Melo, M. Liu, D.L. Neuhoff, "On the Many-to-One Transport Capacity of a Dense Wireless Sensor Network and the Compressibility of Its Data", *Proc. Inter. Conference on Information Processing in Sensor Networks (IPSN)*, LNCS 2634, pp. 1–16, 2003.
- [13] M.J. Neely, E. Modiano, "Capacity and Delay Tradeoffs for Ad Hoc Mobile Networks", *IEEE Trans. on Information Theory*, vol. 51, n. 6, pp. 1917–1937, 2005.
- [14] N. Priyantha, H. Balakrishnan, E. Demaine, S. Teller, "Anchor-Free Distributed Localization in Sensor Networks", *Proc. ACM SenSys*, 2003.
- [15] G. Resta, P. Santi, "Latency and Capacity Optimal Broadcasting in Wireless Multi-Hop Networks", *Proc. IEEE International Conference on Communications (ICC)*, pp. 1–6, 2009.
- [16] G. Resta, P. Santi, "Latency and Capacity Optimal Broadcasting in Wireless Multihop Networks with Arbitrary Number of Sources", *submitted for publication*. Available as Tech. Rep. IIT-TR02/2009, Istituto di Informatica e Telematica, Pisa, Italy.
- [17] G. Resta, P. Santi, "On the Fundamental Limits of Broadcasting in Wireless Mobile Networks", *Proc. IEEE Infocom (miniconference)*, to appear, 2010.
- [18] C. Scheideler, A. Richa, P. Santi, "An  $O(\log n)$  Dominating Set Protocol for Wireless Ad Hoc Networks under the Physical Interference Model", *Proc. ACM MobiHoc*, pp. 91–100, 2008.
- [19] S. Shakkottai, X. Liu, R. Srikant, "The Multicast Capacity of Large Multihop Wireless Networks", *Proc. ACM MobiHoc*, pp. 247–255, 2007.
- [20] G. Sharma, R. Mazumdar, N. Shroff, "Delay and Capacity Trade-offs in Mobile Ad Hoc Networks: A Global Perspective", *Proc. IEEE Infocom*, pp. 1–12, 2006.
- [21] F. Sivrikaya, B. Yener, "Time Synchronization in Sensor Networks: a Survey", *IEEE Network*, Vol. 18, n. 4, pp. 45–50, 2004.
- [22] G. Xing, X. Wang, Y. Zhang, C. Lu, R. Pless, C. Gill, "Integrated coverage and connectivity configuration for energy conservation in sensor networks", *ACM Transactions on Sensor Networks*, Vol. 1, n. 1, pp. 36–72, 2005.
- [23] R. Zheng, "Information Dissemination in Power-Constrained Wireless Networks", *Proc. IEEE Infocom*, pp. 1–10, 2006.
- [24] R. Zheng, R. Barton, "Towards Optimal Data Aggregation in Random Wireless Networks", *Proc. IEEE Infocom*, pp. 249–257, 2007.



**Paolo Santi** received the Laura Degree *summa cum laude* and Ph.D. Degree in Computer Science from the University of Pisa in 1994 and 2000, respectively. He has been researcher at the Istituto di Informatica e Telematica - CNR in Pisa, Italy, since 2001, and Senior Researcher since 2009. During his career, he visited Georgia Institute of Technology in 2001, and Carnegie Mellon University in 2003. His research interests include fault-tolerant computing in multi-processor systems (during Ph.D. studies), and, more recently, the investigation of fundamental properties

of wireless multihop networks such as connectivity, topology control, lifetime, capacity, mobility modeling, and cooperation issues. He has contributed more than 50 papers and a book in the field of wireless ad hoc and sensor networking, he has been General Co-Chair of ACM VANET 2007/2008, Technical Program Co-Chair of IEEE WiMesh 2009, and he is involved in the organizational and technical program committee of several conferences in the field. Since February 2008, Dr. Santi is Associate Editor for IEEE Transactions on Mobile Computing. He is a senior member of ACM and SIGMOBILE.