On The Scaling Laws of Dense Wireless Sensor Networks

Hesham El Gamal*

April 12, 2003

Abstract

We consider dense wireless sensor networks deployed to observe arbitrary random fields. The requirement is to reconstruct an estimate of the random field at a certain *collector node*. This creates a *many-to-one* data gathering wireless channel. In this paper, we first characterize the transport capacity of many-to-one dense wireless networks subject to a constraint on the total average power. In particular, we show that the transport capacity scales as $\Theta(\log(N))$ when the number of sensors N grows to infinity and the total average power remains fixed. To prove the achievability of this result, we devise a simple constructive approach for realizing this capacity based on an *antenna sharing* idea. Interestingly, this constructive approach shows that one can achieve a transmission rate of the same order as the transport capacity with single user receivers. The transport capacity result is independent of where the traffic is generated in the network. Moreover, we show that the same scaling law holds even if every node has a constant non-vanishing average power. We then use this result along with some information theoretic tools to derive sufficient and necessary conditions that characterize the set of *observable* random fields by dense sensor networks. In particular, for random fields that can be modelled as discrete random sequences, we derive a certain form of source/channel coding separation theorem. We further show that one can achieve any desired *non-zero* mean square estimation error for continuous, Gaussian, and *spatially bandlimited* fields through a scheme composed of single dimensional quantization, distributed Slepian-Wolf source coding, and the proposed antenna sharing strategy. Based on our results, we revisit earlier conclusions about the feasibility of dense sensor networks. As argued in the sequel, our results may have some important implications on the design of dense sensor networks.

*The author is with the Electrical Engineering Department, the Ohio State University. This work was supported by the National Science Foundation under a small ITR grant. Parts of this work have been submitted to VTC'03 and SenSys'03.

1 Introduction

In a seminal paper, Gupta and Kumar have shown that the capacity of large scale *ad-hoc* wireless networks scales as $\Theta\left(\sqrt{N}\right)$ as the number of nodes N per unit area grows to infinity [1]. This result means that the capacity per node only scales as $\Theta\left(\sqrt{\frac{1}{N}}\right)$, and hence, goes to zero as $N \to \infty$. While this result advises against deploying dense *ad-hoc* networks, the situation may be different in the context of wireless sensor networks. The enabling observation in this context is that the traffic generated at the different sensors is not *independent* as in the case of ad-hoc networks studied by Gupta and Kumar [2, 3]. In fact, the correlation between the traffic generated at adjacent sensor nodes increases as the density of the sensor nodes per unit area grows. Therefore, as $N \to \infty$, the high correlation between the different observations will result in a traffic *per sensor node* that goes to zero [2, 3].

There is, however, another important difference between the sensing application and the model used by Gupta and Kumar which was recently observed in [4]. The scenario investigated in [1] assumes *peer-to-peer* communication where the information generated at an *arbitrary* node is transported to another *arbitrary* node. The sensing application, however, implies a fundamental difference in the topology of the network from this peer-to-peer scenario. In this paper, we focus on the case where all the sensing information must be collected at *a single* node. We will refer to this scenario as the *many-to-one* channel and to the information sink as the collector node. This architectural difference in the network topology has implications on the maximum traffic carrying capacity of the network. For example, in [4] the authors used the modelling assumptions of [1] to derive the following upper bound on the capacity of this many-to-one channel: Based on the assumption that every node can transmit or receive a maximum of W bits per second [1], it is straightforward to see that the capacity of the many-to-one channel is upper bounded by W. The first contribution of this paper is to show that this result is, in fact, *over-restrictive*.

In particular, we show that the transport capacity scales as $\Theta(\log(N))$ when the number of sensors N grows to infinity and the total average power remains fixed. An important part of our approach is a novel transmission scheme that exploits the high density of sensor nodes to facilitate *antenna sharing* at minimal cost in resources. We show that this scheme achieves the same scaling law as the optimal approach while employing single user receivers (i.e., receivers that attempt to detect only one information stream at any point in time treating all other information streams as noise). One of the interesting insights allowed by our proof is that, contrary to the *peer-to-peer* scenario, spatial reuse of the bandwidth does not factor prominently in this many-to-one case. In a nutshell, one san say that the *single sink node* in the many-to-one channel acts like a *bottleneck*, as predicted by [4]. The cost entailed by this bottleneck is, however, not as dramatic as argued in

[4] (instead of lowering the traffic from $\Theta\left(\sqrt{N}\right)$ to $\Theta(1)$, we show that the traffic is only reduced to $\Theta(\log(N))$).

We then use this result to characterize the set of *observable*[1] random processes by this class of dense sensor networks. For random fields that can be modelled as discrete random sequences, we derive necessary and sufficient conditions for observability. We further establish a certain form of source/channel coding separation theorem in this scenario. The significance of this separation theorem is that it allows for constructing low complexity coding and decoding algorithms. Next, we investigate the more realistic, and challenging, scenario of continuous random processes. In this case, we show that all Gaussian and spatially bandlimited processes can be estimated at the collector node, subject to any non-zero constraint on the mean square error, using a simple strategy composed of single dimensional quantization, Slepian-Wolf distributed source coding, and the proposed antenna sharing approach. Towards the end of the paper, we shed some light on the limitation of the multiple access formulation used earlier to investigate the capacity of the many-to-one channel [5].

The rest of this paper is organized as follows. In Section 2, we present our modelling assumptions and notation. The main results regarding the transport capacity of the many-to-one wireless channel is developed in Section 3. In Section 4, we report our results regarding the observability of discrete and continuous spatial random processes by dense wireless networks. Section 5 sheds some light on the limitations of the multiple access formulation when used to study the many-to-one wireless channel. Finally, we offer some concluding remarks in Section 6.

2 System Model and Assumptions

For simplicity of presentation, unless otherwise stated, we consider the scenario where the N sensor nodes are distributed uniformly over the perimeter of a circle with a unit radius. The collector node is assumed to be at the center of the circle, and hence, all the source nodes are at the same distance from the collector. We will, however, show later that our results hold for arbitrary single dimensional networks[2]. It will become clear in the sequel that most of our results extend to the planar scenario where the sensor nodes occupy the surface of a plane. Moreover, for large N, the results hold with high probability for a random, but a uniform, distribution of the sensors. We will refer to the distance between the i^{th} and j^{th} sources as $d_{i,j}$, where $i, j \in \{1, ..., N\}$. We reserve the index "0" for the collector node, and hence, the distance between the i^{th} source and the collector is $d_{i,0} = 1$. We further assume that the source nodes are equipped with receivers. Our result holds

[1]This notion of observability will be rigorously defined in the sequel.

[2]The sensor nodes are distributed uniformly over an arbitrary single dimensional curve

whether the source nodes can transmit and receive simultaneously or can only do one task at a time (i.e., transmit or receive). We assume that all nodes share a wireless medium with a finite bandwidth. We use a discrete time model where $y_j[k]$, the received signal by the j^{th} node at time k, is given by

$$y_j[k] = \sum_{i \in \{1,...,N\}, i \neq j} \frac{e^{j\theta_{i,j}} x_i[k]}{d_{i,j}^{\delta}} + n_j[k], \quad (1)$$

$x_i[k]$ is the signal transmitted by node i at time k, δ is the path loss exponent assumed to be strictly larger than zero [6], $\theta_{i,j}$ is the phase shift resulting from the propagation delay, and $n_j[k]$ is the zero-mean and unit variance additive Gaussian noise sample at receiver j and time k. The noise samples are assumed to be spatially and temporally independent. The phase shift between the i^{th} and j^{th} nodes, i.e., $\theta_{i,j}$, is assumed to be known at both nodes. In practice, these parameters can be estimated at a marginal loss in throughput (the loss goes to zero as the time scale of the network operation goes to infinity). In (1), it is assumed that all the sensor nodes are synchronized with a common clock. We further assume that the network operates in slotted frames where the duration of one slot T_s is long enough to allow for invoking the asymptotic additive white Gaussian noise (AWGN) channel capacity theorem. Without loss of generality, we focus our analysis on an arbitrary time slot, and hence, assume that $k \in \{1, ..., T_s\}$. The path loss model used in (1) implicitly assumes that all the nodes, including the collector node, use identical omnidirectional antennas. We denote the observed random variable by node j at time k as $u_j[k]$. For the class of continuous random processes, we only consider temporally stationary and bandlimited processes. We further assume that the process is sampled at the Nyquist rate such that $u_j[k_1]$ and $u_j[k_2]$ are independent and identically distributed for any $k_1 \neq k_2$ and arbitrary j. The same i.i.d assumption is adopted for discrete processes as well. The spatial observations $u_j[k]$ and $u_i[k]$ are, however, correlated. In fact, the spatial correlation between observations at adjacent nodes is expected to grow as the density of sensors increases. We assume that the joint distribution of all the observations is known *a-priori* at all sensor nodes. This assumption facilitates distributed source coding in the Slepian-Wolf sense as discussed in the sequel. In general, the transmission time of one symbol can be different from the time that transpires between taking two consequent samples $u_j[k]$ and $u_j[k+1]$. Thus, we assume that during a slot of T_s *transmission symbols*, we collect T_u observations at every sensor node. The ratio between T_u and T_s determines the ratio between the transmission bandwidth and the *temporal* bandwidth[3] of the observed process. In this work, we neglect issues related to transmission and processing delays.

[3]We use this notation to differentiate between the temporal and spatial bandwidths of the observed process.

Compared to the system model in [4], we have two *fundamental* differences in our modelling assumptions. It was assumed in [4] that there is a number W such that each sensor can transmit or receive at most W bits in one slot. We substitute this assumption with a constraint on the total average power consumed by the network, i.e.,

$$\frac{1}{T_s}\sum_{k=1}^{T_s}\sum_{i,\ldots,N}|x_i[k]|^2 \leq P_{total}, \tag{2}$$

where P_{total} is the total average power assigned to the network[4] which is finite and for technical convenience is assumed to be $\geq 2\pi$ (this assumption, however, does not affect the conclusions for networks with large N). Together with the finite bandwidth of the shared wireless medium, we believe that this is a more faithful representation of the wireless channel constraints. We also relax the assumption that every node can only receive from one source node at a time. The constructive approach used in our proof, however, shows that one can achieve the same order of transport capacity by using single user receivers (i.e., every node only decodes one information stream at a time treating all other transmissions as noise).

3 The Transport Capacity of the Many-to-One Channel

In this section, we assume that the information streams generated at the different nodes are independent. The implications of the correlation between the observations are investigated in the next section. In this context, the transport capacity of the many-to-one channel C_N is defined as the maximum number of bits that can be transported from the N source nodes to the collector per unit time (in our terminology, unit time refers to the duration of transmission of a single symbol). We say that $C_N = \Theta(\log(N))$ if there are positive constants c_1, c_2, c_3, c_4 (c_2, c_4 are allowed to be zero) such that

$$c_1 \log(N) + c_2 \leq C_N \leq c_3 \log(N) + c_4. \tag{3}$$

log(.) is assumed to have a base 2. Now, we are ready to prove the following result that characterizes the scaling law of the transport capacity in dense sensor networks.

Theorem 1 *The transport capacity of the many-to-one channel outlined in Section 2 is* $C_N = \Theta(\log(N))$.

Proof: The upper bound on the transport capacity of the many-to-one channel is easily obtained by considering the scenario where a genie informs every source node of the information generated at

[4] P_{total} is normalized to refer to the total received average power at a unit distance from the transmitter.

all other source nodes. In this case, the many-to-one channel is equivalent to a one-to-one channel with N transmit antennas and one receive antenna. The optimal solution in this scenario is to assign equal power to every antenna and use a Gaussian codebook followed by a beamformer (i.e., the transmitted signal from the i^{th} node is multiplied by $e^{-j\theta_{i,0}}$ to allow for coherent combining at the collector node). The capacity of this multi-antenna one-to-one channel is given by (e.g., [7])

$$\begin{aligned} C_N &= \log\left(1 + NP_{total}\right) \\ &\leq \log\left(2P_{total}N\right) \\ &= c_3 \log\left(N\right) + c_4, \end{aligned} \tag{4}$$

where $c_3 = 1$ and $c_4 = \log\left(2P_{total}\right)$.

Establishing the lower bound is more challenging. To this end, we devise a simple transmission protocol that exploits the high density of the nodes to facilitate cooperative transmission. For simplicity of presentation, we consider the symmetric scenario where all the source nodes generate the same amount of traffic. It is shown later, however, that the result can be extended to arbitrary asymmetric scenarios through a minor modification of the transmission protocol.

The main idea in our protocol is to allow every node to distribute its information to closely located nodes, which comes at a very small cost for densely deployed networks, and then those nodes can cooperate to transmit the information to the collector using a beamformer to get the logarithmic increase in the **received** power, and subsequently, the capacity. In our protocol, every node transmits its information to the collector in two time slots. Without loss of generality, we consider the transmission of the information stream from node "1". In the first time slot, P_{total} is assigned to node "1" and all the other nodes are only listening. The transmission occurs at a rate of $\alpha \log(N)$ per unit time in this slot (in the following we will obtain α and show that it is > 0). An arbitrary node j will be able to decode correctly if and only if $C_{1,j} \geq \alpha \log(N)$, where

$$C_{1,j} = \log\left(1 + \frac{P_{total}}{d_{1,j}^{2\delta}}\right). \tag{5}$$

Now, we find a lower bound on the nodes that can decode correctly. We let $d_{1,j} = N^{-\gamma_{1,j}}$, where $\gamma_{1,j}$ is a positive real number. Then a sufficient condition for the j^{th} node to successfully decode the transmission is

$$\log\left(N^{2\gamma_{i,j}\delta}\right) \geq \alpha \log(N). \tag{6}$$

or equivalently

$$2\gamma_{i,j}\delta \geq \alpha. \tag{7}$$

So, all the nodes within a distance of $N^{-\frac{\alpha}{2\delta}}$ will be able to decode successfully. Since the distance between two neighbors on the circle is upper bounded by $\frac{2\pi}{N}$ due to the uniform distribution of the source nodes, the number of nodes successfully decoding the transmission is readily lower bounded by $\frac{1}{2\pi}N^{\left(1-\frac{\alpha}{2\delta}\right)}$. In the second time slot all the $\frac{1}{2\pi}N^{\left(1-\frac{\alpha}{2\delta}\right)}$ nodes cooperate with node 1, in a beamforming configuration with equal transmit power assigned to every node, to deliver the information to the collector node. All other nodes are silent in this slot. A sufficient condition for the collector node to decode successfully is[5]

$$\log\left(N^{\left(1-\frac{\alpha}{2\delta}\right)}\right) \geq \alpha\log(N). \tag{8}$$

This condition is satisfied if α is chosen to be

$$\alpha = \frac{2\delta}{2\delta+1}. \tag{9}$$

Since it took us two time slots to deliver $\frac{2\delta}{2\delta+1}T_s\log(N)$ bits to the collector, the throughput of the proposed protocol is lower bounded by

$$C_N \geq c_1\log(N), \tag{10}$$

where $c_1 = \frac{\delta}{2\delta+1} > 0$. Combining (4) and (10), we obtain our result.

$$C_N = \Theta\left(\log(N)\right). \tag{11}$$

The final step is to symmetrize the transmission protocol by assigning every two consecutive time slots to a different source node. □

Proposition 2 *The transport capacity result in Theorem 1 holds if the sensor nodes are distributed uniformly over an arbitrary curve of length $\leq 2\pi$, and for an arbitrary location of the collector node satisfying the constraint the maximum distance between the collector node and any senor node $d_{max} \leq 1$.*

Proof: First, assume that the collector node is not located at the same curve as the sensors. Let $d_{min} \leq 1$ be the minimum distance between the collector node and any point on the line where the sensor nodes are located. It is easy to see that d_{min} does not change as the number of sensor grows. Then the following upper bound on the transport is obtained by adopting the genie aided strategy

[5]Here, we need the technical assumption of $P_{total} \geq 2\pi$ so that our result is valid for any N.

along with the assumption that all the sensors are at the same distance, d_{min}, from the collector node

$$\begin{aligned} C_N &\leq \log\left(1 + \frac{P_{total} N}{d_{min}^{2\delta}}\right) \\ &\leq c_3 \log(N) + c_4, \end{aligned} \tag{12}$$

where $c_3 = 1$, $c_4 = \log(\frac{2P_{total}}{d_{min}^{2\delta}}) \geq 0$. The achievability part can be easily proved by assuming that all the sensor nodes are at the same distance $d_{max} \leq 1$ from the collector and using the constructive approach outlined in the proof of Theorem 1.

If the collector node is located on the same curve as the sensor nodes then the only difference is that the minimum distance between the collector node and the sensors will now decrease as $1/N$ as the number of sensors grows. This clearly does not affect the achievability part of the result since the constraint on d_{max} remains the same. One can further see that the upper bound in (12) in this case still holds, however, with a different $c_3 = 1 + 2\delta$. This completes the proof.

□

A few remarks about the transport capacity result are now in order. In particular, we show how Theorem 1 allows for different conclusions than those drawn in [1, 6, 4] and highlight the reasons behind these differences.

1. Theorem 1 and Proposition 2 are valid for any N not necessarily large. In order to facilitate this generality, we needed the technical assumption of $P_{total} \geq 2\pi$, $d_{max} \leq 1$, and the length of the curve where the sensors are located is $\leq 2\pi$. However, we realize these results become only significant for large N since for small values of N the performance of the network may be dominated by the constants c_1, c_2, c_3, c_4 rather than the $\log(N)$ term. For large N, the need for these technical assumptions disappears since one can now allow c_2 to be negative without affecting the conclusions. The only necessary restriction is that the length of the curve does not grow as the number of sensors grows. Also, in this asymptotic scenario, one can see that the results hold with high probability if the positions of source nodes are chosen according to a uniform i.i.d assumption, rather than uniformly, on the curve. Finally, it is straightforward to extend these results to planar networks where the sensor nodes cover the surface of a two dimensional plane.

2. It is interesting to note that although we have allowed for simultaneous transmissions from multiple nodes in the proposed antenna sharing approach, this *does not* entail the use of any

sophisticated multi-user detection schemes. The reason is that all the transmitting nodes are cooperating to deliver the same information stream.

3. One of the significant implications of Theorem 1 is that one can achieve an *unbounded* transport capacity for the many-to-one channel as the number of nodes grows to infinity with only *finite* total average power. More interestingly, even if one allows for the total power to scale linearly with the number of nodes N, we will get the same scaling law for the transport capacity (i.e., $\Theta(\log(N))$. This is easily verified from the beamforming upper bound. One can attribute this limitation to the *bottleneck* resulting from the need to deliver all the information to a single collector node. In summary, the *magic* $\Theta\left(\sqrt{N}\right)$ in the peer-to-peer scenario translates into $\Theta(\log(N))$ in the many-to-one scenario.

4. Theorem 1 was proven for the symmetric scenario where the traffic generated at each node is the same. It is, however, straightforward to extend it to arbitrary asymmetric scenarios through a minor modification of the communication protocol where every node is assigned a number of time slots proportional to the amount of traffic generated at this node.

5. In [4], Marco *et al.* investigated the feasibility of the data gathering channel in dense sensor networks. There, the authors argued that the transport capacity is upper bounded by a constant and does not scale as the number of sensors N grows to infinity. This conclusion is based on the assumption that every node can receive at most W bits per second which creates a bottleneck at the collector node. We believe that this assumption *does not* correspond to specific constraints on the *resources* available to the wireless network in terms of average power, bandwidth, or receiver complexity. In fact, the same assumption results in a finite total average power consumed in the many-to-one channel [4] and an infinite total power consumed in the peer-to-peer scenario [1]. Here, we substitute this assumption with two constraints, that the bandwidth of the shared wireless medium is finite and the total average power is finite, and a modelling assumption on the path loss. With this set-up, we showed that the transport capacity scales as $\log(N)$ for a dense network with N sensors. In summary, one can attribute the difference in conclusions between our work and [4] to the fact that the assumption used in [4] does not allow for exploiting the high density of sensors to facilitate efficient cooperative transmission strategies.

6. The results in [1] indicate that spatial reuse of the available resources is one of the key components that should be exploited in wireless networking. Quite interestingly, spatial reuse does not seem to factor prominently in the many-to-one scenario considered here. This is evident in the constructive approach used to prove the achievability part of Theorem 1 where

no spatial reuse is allowed (i.e., all the resources are dedicated to transporting one information stream to the collector node). One can attribute this difference to the architectural difference between the peer-to-peer scenario considered in [1] and the current scenario. In our scenario, all the sources are *competing* not only for the wireless medium but also for the same *destination*. It is clear that spatial reuse will not help in resolving the competition for the same destination which is the dominant factor that dictates the $\Theta(\log(N))$ scaling law.

7. In [6], Xie and Kumar concluded that the information theoretic capacity of wireless networks is upper bounded by the total average power assigned to the network, even if the number of nodes $N \to \infty$. The seeming contradiction between our conclusion and that in [6] can be traced back to one fundamental difference in the network topology. In [6], the minimum distance between two nodes is lower bounded by a certain value ρ_{min}, and hence, increasing the number of nodes *necessarily* corresponds to an increase in the geographical coverage of the network. Here, we restricted the geographical size of the network to be independent of the number of nodes, and hence, increasing the number of nodes translates into a smaller distance between adjacent nodes. We believe that our assumption is more representative of the sensing application where increasing N should translate into a larger density of sensors per unit area.

4 The Observability of Spatial Random Processes

Now, we investigate the effect of the correlation between the observations at the different sensors on the operation of wireless sensor networks. We show that the transport capacity result in Theorem 1 still plays a significant role in this scenario. Specifically, we utilize Theorem 1 along with some information theoretic tools to develop necessary and sufficient conditions on the observability of random processes by dense wireless sensor networks. Moreover, we establish some interesting results regarding low complexity schemes that allow for the separation between source and channel coding. Throughout this section, we focus on the asymptotic scenario $N \to \infty$. Our definition of observability and our results depend largely on whether the observed random variables are discrete or continuous. Before proceeding further, we need the following definition.

Definition 3 *We say that $f(N) \leq \Theta(\log(N))$ if as $N \to \infty$ there exists a positive and finite c such that $f(N) \leq c\log(N)$. Similarly, we say that $f(N) > \Theta(\log(N))$ if $f(N) = c\log(N)$ implies that $c \to \infty$ as $N \to \infty$.*

4.1 Discrete Sources

In this scenario, the random variables $u_j[k]$ are assumed to be discrete. The discreteness of the observed random variables implies that one should attempt to reconstruct them with arbitrarily small probability of error[6] at the collector node. This is reflected in the following definition of *observability*

Definition 4 *A discrete random sequence is said to be observable if it can be* **detected** *at the collector node with arbitrarily small probability of error for a certain allocation of a finite bandwidth and a finite total average power to the network.*

Theorem 5 *The spatial random sequence is observable in the sense of Definition 4 if and only if*

$$H\left(u_1[k], u_2[k], ..., u_N[k]\right) \leq \Theta\left(\log(N)\right),$$

where $H(.,.,.)$ refers to the joint entropy. Moreover, this result holds even if we restrict ourselves to schemes where source and channel coding are performed separately (i.e., source/channel coding separation).

Proof: The converse follows from the *genie aided* strategy where all the sensors are informed of the observations of the other sensors. This transforms the many-to-one wireless channel to a one-to-one channel with N transmit antennas. One can now use the classical source/channel coding separation result for the one-to-one channel to show that if

$$H\left(u_1[k], ..., u_N[k]\right) > \frac{T_s}{T_u} \log\left(1 + P_{total} N\right), \tag{13}$$

for every finite T_s/T_u and P_{total}, then the collector node cannot observe this sequence with arbitrarily small probability of error for any choice of finite bandwidth and finite total average power.

In order to prove the achievablity part of the Theorem we follow a two step approach. This two step approach will also serve to prove the source/channel coding separation result. In the first step, the spatial random sequence is compressed using a Slepian-Wolf distributed source coding algorithm. In particular, the observation vector at the first sensor, i.e., $[u_1[1], ..., u_1[T_u]]$ is compressed to a rate $T_u/T_s H(u_1[k])$ bits per symbol. The second source is then compressed to a rate $T_u/T_s H(u_2[k]|u_1[k])$ bits/symbol. The procedure is then continued such that the j^{th} source is compressed to a rate $T_u/T_s H(u_j[k]|u_1[k], ..., u_{j-1}[k])$ bits per symbol. It is then easy to see that the total rate generated

[6] We use the standard definition of probability of error as in [8], for example.

by the N sensors is $T_u/T_s H(u_1[k], ..., u_N[k])$ bits per symbol. We find it appropriate to stress here that in this procedure every sensor node only needs to know that the joint distribution of the spatial observations (and not the actual observations made at the other sensors). In the second step, we use the cooperative transmission strategy developed in the proof of Theorem 1. As shown in Theorem 1, this transmission strategy supports a rate C_N that satisfies the following lower bound

$$C_N \geq \frac{\delta}{2\delta+1} \log\left(N\right),$$

independent of where the traffic is generated in the network. So, if

$$H\left(u_1[k], ..., u_N[k]\right) \leq \log(N),$$

then we can choose a finite transmission bandwidth (to adjust T_u/T_s) such that

$$\frac{T_u}{T_s} H\left(u_1[k], ..., u_N[k]\right) = \frac{\delta}{2\delta+1} \log\left(N\right).$$

This establishes the achievability of the Theorem. □

Here, we remark that our separation result only establishes that one can achieve the same scaling law as the optimal scheme with a two step approach. This does not, necessarily, guarantee that the two step approach will achieve the same throughput as the optimal scheme (i.e., there may be a difference in the constants).

4.2 Continuous Sources

The scenario of continuous sources is more realistic than its discrete counterpart and also more challenging. The additional challenge here stems from the fact that the distributed *lossy*[7] source coding problem is still open [8]. Here, we use the simple approach proposed in [4] for distributed lossy source coding. This approach is composed of single dimensional quantization followed by distributed Slepian-Wolf coding for the resulting discrete valued random sequence. We denote the quantized discrete random variable at sensor j and time k as $v_j[k]$. Our results in this section are limited to single dimensional sensor networks observing spatially stationary random processes. The

[7]With finite communication resources, one can only reproduce a single continuous random variable with finite non-zero distortion. Hence, one must resort to lossy source coding techniques in the scenarios involving continuous random variables.

spatial stationarity implies that the process is isotropic. Before proceeding further, we need the following definitions.

Definition 6 *The spatial stationarity assumption allows for parameterizing the spatial correlation as a function of the distance between the random variables, and hence, the correlation between* $u_i[k]$ *and* $u_j[k]$ *is referred to as* $R(d_{i,j})$. *A random process is said to be spatially bandlimited if the the Fourier transform of* $R(d)$, $S_R(f)$, *satisfies the property that* $S_R(f) = 0$ *if* $|f| \geq f_0$ *for a an arbitrary* $f_0 < \infty$. *We therefore refer to* f_0 *as the spatial bandwidth of the process.*

Definition 7 *A continuous random process is said to be observable if it can be* **estimated** *at the collector node with a finite non-zero mean square error for a certain allocation of a finite bandwidth and a finite total average power to the network.*

We are now ready to prove our main result in this section.

Theorem 8 *All Gaussian spatially bandlimited processes are observable by dense wireless sensor networks as* $N \to \infty$. *Moreover, this result holds even if we restrict ourselves to the scheme composed of a lossy source encoder (single dimensional quantization followed by Slepian-Wolf distributed coding) concatenated with the proposed cooperative transmission strategy (i.e., source/channel coding separation).*

Proof: Using the result we have established earlier for discrete sources, it suffices now to show that the joint entropy of the quantized random variables satisfies the following upper bound

$$H\left(v_1[k], ..., v_N[k]\right) \leq \Theta\left(\log(N)\right) \tag{14}$$

To show this, we first need the following upper bound from [4]

$$\begin{aligned} H\left(v_1[k], ..., v_N[k]\right) &= H(v_1[k]) + \sum_{i=2}^{N} H\left(v_i[k]|v_{i-1}[k], ..., v_1[k]\right) \\ &\leq H(v_1[k]) + \sum_{i=2}^{N} H\left(v_i[k]|v_{i-1}[k]\right) \\ &= H(v_1[k]) + (N-1)H\left(v_2[k]|v_1[k]\right) \\ &\to NH\left(v_2[k]|v_1[k]\right), \end{aligned} \tag{15}$$

where the first step follows from the chain rule, the second step from the fact that conditioning reduces entropy, the third step from the spatial symmetry of the process, and the fourth step from taking the limit as $N \rightarrow \infty$. We will also need the following result from [4] for Gaussian random processes

$$\lim_{R(d_{1,2})\rightarrow 1} \frac{H\left(v_2[k]|v_1[k]\right)}{-\sqrt{1-R^2(d_{1,2})}\log\left(c\sqrt{1-R^2(d_{1,2})}\right)} = 1, \tag{16}$$

where c is a constant that depends on the quantization step. This result is useful since as $N \rightarrow \infty$ one would expect $R(d_{1,2}) \rightarrow 1$. The final ingredient is to characterize the behavior of $R(d_{1,2})$ as $N \rightarrow \infty$ for bandlimited processes. First, we know that $R(d_{1,2}) = R(a/N)$ since $d_{1,2} \leq 2\pi/N$. The bandlimited nature of the process means that

$$R(a/N) = \int_{-f_0}^{f_0} S_R(f)\cos(2a\pi f/N)df. \tag{17}$$

For large N, one can use the Taylor series expansion of cos(.) to show that

$$R(a/N) \approx 1 - \frac{b}{N^2}, \tag{18}$$

where $b = 4a^2\pi^2 \int_{-f_0}^{f_0} f^2 S_R(f)df < \infty$. Then

$$R^2(a/N) \approx 1 - \frac{2b}{N^2}. \tag{19}$$

One can now use (16) and (19) to obtain

$$H(v_2[k]|v_1[k]) \approx \frac{\sqrt{2b}}{N}\log\left(\frac{cN}{\sqrt{2b}}\right) \tag{20}$$

.

Using (15), we then obtain the desired result

$$H(v_1[k], ..., v_N[k]) \leq \Theta\left(\log(N)\right). \tag{21}$$

□

A few remarks are now in order

1. While Theorem 8 is stated for bandlimited processes, clearly the result extends to other processes that *generate* $\leq \Theta\left(\log(N)\right)$ bits per transmission symbol. The Gaussian process with $R(d) = e^{-d^2}$ given in [4] is one example of such processes.

2. Theorem 8 assumes that a particular source coding scheme of single dimensional quantization at all the sensors followed by Slepian-Wolf distributed source coding for the quantized process. We don't claim any optimality properties for this scheme. For example, with bandlimited processes, the total number of bits generated by the sensors can be lowered if we only activate the sensors located at the spatial Nyquist intervals for observations. The remaining sensors can be still used for cooperative transmission. This approach, however, requires a dynamic activation strategy that chooses the active sensors based on the spatial bandwidth of the observed process which may be time varying. The scheme considered in Theorem 8 is, therefore, more robust to these temporal variations since all the sensors are *actively* observing the process always.

3. Distributed source coding for sensor networks observing continuous random processes was considered in [9]. There, the author presented an example for a feasible dense wireless sensor network which still generates a traffic of the order $\Theta\left(\sqrt{N}\right)$. The seeming contradiction between [9] and our conclusion that the traffic of a feasible sensor network has to be in the order of $\Theta\left(\log(N)\right)$ can be attributed to the network architecture considered in [9]. This network architecture consists of a grid of nodes where the nodes on the left vertical line are the sensor nodes observing the random process (source nodes), the nodes on the right vertical line are the sinks, and the nodes in the middle only act as relays. Every source on the left transmits its information *only* to the sink on the right side connected to it via a horizontal line. Then, the author assumed that the collection of the *correlated* information from the sink nodes to the central collector is *free of charge*. This assumption is the key behind the difference in conclusions. Neglecting the cost of information collection in the last step implies the assumption that we have a collector node with infinite number of receive antennas (i.e., all the sink nodes on the right are acting as antennas for the collector node). This assumption is significantly different from the scenario considered here where the collector node has only one receive antenna and is assumed to be at *a non-zero* distance from the nearest sensor node.

5 The Limitation of the Multiple-Access Channel Formulation

Earlier works on the analysis of the wireless many-to-one channel have been largely inspired by the multiple access channel formulation [5, 4]. In this section, we revisit this formulation. In doing so, our main goal is to highlight the fact that the wireless channel offers some opportunities that, sometimes, lay beyond the grasp of traditional information theoretic models. While we are clearly not the first to make this observation, we believe that our results offer a clear example that supports this view. The discussion further serves to highlight the significance of the source/channel coding separation theorem established in Section 4.1.

In [5], Barros and Servetto considered the multiple access channel with discrete sources under the assumption that the channels from the different sources to the destination are orthogonal. Based on this assumption, the authors derived the capacity region and further established the source/channel separation theorem for this orthogonal multiple access channel. One can now apply these results to the current additive white Gaussian noise case by adapting a time division multiple access (TDMA) strategy[8] [5]. We assume that the j_1 source node is assigned a fraction f_{j_1} of the time slots. During these slots, the total average power P_{total} is assigned to the j_1 source node and all the other $N-1$ are silent. We can now write the capacity region as[9]

$$H\left(u_{j_1}, ..., u_{j_i} | u_{j_{i+1}}, ..., u_{j_N}\right) \leq \frac{T_s}{T_u}(f_{j_1} + ... + f_{j_i})\log(1 + P_{total}), \tag{22}$$

where (22) must be satisfied for any choices of $1 \leq i \leq N$ and $j_1, ..., j_N$ such that the set $\{j_1, ..., j_N\}$ is a permutation of the set $\{1, ..., N\}$. Now, since one has the freedom to assign the fractions $f_1, ..., f_N$ based on the source statistics to maximize the capacity region, it is easy to see that the condition of observability boils down to

$$H\left(u_1, ..., u_N\right) \leq \frac{T_s}{T_u}\log(1 + P_{total}). \tag{23}$$

Furthermore the capacity in (23) is achieved by a Slepian-Wolf distributed source coding algorithm followed by a bank of *independent* channel codes optimized for the AWGN channel (i.e., source/channel coding separation) [5]. One can now safely conclude that this formulation yields a transport capacity $\Theta\left(1\right)$. In fact, one gets the same result if we substitute the orthogonal channels condition with the restriction that the transmitted signals from the different source nodes are independent (i.e., the well known multiple access channel result) [8]. The explanation for the decrease in the transport capacity (from $\Theta\left(\log(N)\right)$ to $\Theta\left(1\right)$) can be now quoted from [8] *"To maximize capacity one should preserve the correlation between the inputs of the channel. Slepian-Wolf encoding, on the other hand, gets rid of the correlation"*. One possible way to enhance the achievable rate, without abandoning the multiple access channel formulation, is to exploit the correlation between the information streams to transmit *correlated* signals from the different sources [10]. This approach,

[8] The sensor nodes are again assumed to be located on the unit circle with the collector at the center

[9] Here we remove the reference to the time index k for simplicity

however, entails more complexity since no source/channel coding separation can be claimed. Furthermore, finding a closed form formula for the achievable rate using this approach does not seem to be a straightforward task.

In a nutshell, similar to the multiple access channel, the transport capacity of the many-to-one wireless channel is maximized if the transmissions from the different sources are highly correlated. In the multiple access scenario, one can only resort to the correlation between the different sources to create a *statistical* dependence between the different transmitted signals. This approach, therefore, rules out the separation between source and channel coding, and hence, presents a potential risk for high computational complexity. On the other hand, in the many-to-one wireless sensor network, one can exploit the high density of sensors and the wireless environment to facilitate *low cost* inter-sensor communication. This allows for creating a *deterministic* dependence between the transmitted signals (i.e., beamforming). This way, the capacity of the many-to-one channel can be maximized (i.e., achieve the same scaling law as the optimal scheme) while preserving the desirable separation between source and channel coding.

6 Concluding Remarks

In this paper, we first characterized the transport capacity scaling law of the many-to-one wireless channel. In particular, we have established that Gupta and Kumar *magic* $\Theta\left(\sqrt{N}\right)$, which was derived under the peer-to-peer assumption, translates to $\Theta\left(\log(N)\right)$ in the many-to-one scenario. We then introduced the notion of an *observable* random field and used the transport capacity result, along with some information theoretic tools, to develop sufficient and necessary (only in the discrete case) conditions that govern the observability of discrete and continuous random fields. Throughout the paper, we utilized our results to extract, what we believe are, valuable insights in the design of efficient wireless sensor networks. For example, we established the gain offered by the antenna sharing idea, argued for the separation between source and channel coding in certain scenarios, and highlighted the fact that spatial reuse may not be needed in the many-to-one channel.

We believe that our work suggests some interesting venues for future work. Here, we briefly outline two examples of these open problems. The transport capacity result developed here hinges on the model we used for the path loss between the transmitter and receiver. We realize that this model may not be very accurate since the path loss model is based on the far field wave propagation assumption which may not hold when the transmitter and receiver are very close. This concern should motivate revisiting the analysis with more refined models. Even with those refined models, one would still expect that, in dense networks, every node can distribute its information to its neighbors at minimal cost. Hence, we conjecture that the main idea behind the transport capacity result (i.e., utilizing the proximity of source nodes in dense networks to facilitate efficient cooperation protocols) will play a significant role with these models as well.

The second open problem is to characterize the transport capacity of the sensor broadcast problem studied by Scaglione and Servetto [3]. In this scenario, the information generated at **all** the nodes must be broadcasted to **all** other nodes [3]. In [3], the authors used the *peer-to-peer* scaling law developed in [1] as an upper bound on the traffic carrying capacity of the network. It is clear that the transport capacity of the many-to-one channel is a tighter upper bound on the capacity of the all-to-all channel [4]. It is, therefore, reasonable to revisit the conclusions and designs in [3] based on our results.

7 Acknowledgment

The author would like to thank Prof. D. L. Neuhoff for discussions that motivated writing this paper. He would also like to thank Profs. M. Liu and S. Servetto for answering questions related to their papers.

References

[1] P. Gupta and P. R. Kumar. The capcity of wireless networks. *IEEE Transactions on Information Theory*, 46:388–404, March 2000.

[2] D. L. Neuhoff and D. Marco. Distributed encoding of sensor data. *The IEEE Information Theory Workshop, Bangalore, India*, 2002.

[3] A. Scaglione and S. Servetto. On the interdependence of routing and compression in multi-hop sensor networks. *The Proceedings of the 8th ACM International Conference on Mobile Computing and Networking (MobiCom), Atlanta, GA*, 2002.

[4] D. Marco, E. J. Duarte-Melo, M. Liu, and D. L. Neuhoff. On the many-to-one transport capacity of a dense wireless sensor network and the compressibility of its data. *to appear in the International Workshop on Information Processing in Sensor Networks*, 2003.

[5] J. Barros and S. Servetto. On the capacity of the reachback channel in wireless sensor networks. *The Proceedings of the IEEE Workshop on Multimedia Signal Processing (special session on "Signal Processing for Wireless Networks*, Dec. 2002.

[6] L-L. Xie and P.R. Kumar. A network information theory for wireless communication: Scaling laws and optimal operation. *submitted to IEEE Trans. on Info. Theory*, 2002.

[7] E. Teletar. Capacity of multi-antenna gaussian channels. *Technical Report, AT&T-Bell Labs*, June 1995.

[8] T. Cover and J. Thomas. *Elements of Information Theory*. John Wiley Sons, Inc., New York, 1991.

[9] S. Servetto. Quantization with side information: Lattice codes, asymptotics, and applications in wireless networks. *Submitted to the IEEE Transactions on Information Theory*, 2002.

[10] T. Cover, A. El Gamal, and M. Salehi. The capcity of wireless networks. *IEEE Transactions on Information Theory*, 1980.