

# Minimizing Energy Consumption In Large-scale Sensor Networks Through Distributed Data Compression And Hierarchical Aggregation

Seung Jun Baek, Gustavo de Veciana, *Senior Member, IEEE*, and Xun Su, *Member, IEEE*

**Abstract**—In this paper we study how to reduce energy consumption in large-scale sensor networks which systematically sample a spatio-temporal field. We begin by formulating a distributed compression problem subject to aggregation (energy) costs to a single sink. We show that the optimal solution is greedy and based on ordering sensors according to their aggregation costs—typically related to proximity—and, perhaps surprisingly, it is independent of the distribution of data sources. Next we consider a simplified hierarchical model for a sensor network including multiple sinks, compressors/aggregation nodes and sensors. Using a reasonable metric for energy cost, we show that the optimal organization of devices is associated with a Johnson-Mehl tessellation induced by their locations. Drawing on techniques from stochastic geometry, we analyze the energy savings that optimal hierarchies provide relative to previously proposed organizations based on proximity, i.e., associated Voronoi tessellations. Our analysis and simulations show that an optimal organization of aggregation/compression can yield 8-28% energy savings depending on the compression ratio.

**Index Terms**—Sensor networks, distributed data compression, data aggregation, stochastic geometry

## I. INTRODUCTION

**A**N emerging vision for the future is that of a physical (or virtual) world embedded with networked sensors and actuators. The interweaving of local sensing/actuation, communication, and computation with decision-making and control has broad applicability, including: transportation, environmental monitoring/oversight, structural monitoring, health care, and national security. Progress is being made toward increasing the diversity of available sensors, while new technologies permit the flexible deployment of small, inexpensive devices operating on limited battery power and which are interconnected via wireless links [1], [2], [3], [4]. In this paper we consider distributed compression and aggregation schemes for large scale sensor networks gathering information on a spatio-temporal field. A key challenge for such applications lies in devising system architectures and protocols to realize the required sensing and communication tasks subject to hard system constraints, in particular, energy.

The focus of this paper is on achieving energy efficiency in two ways. First, by simply reducing the traffic transported

by the network, e.g., through distributed source coding[5], [6], [7] and/or data aggregation/header compression[8], [9]. Second, by making the transport of traffic on a sensor network energy efficient, e.g., through energy-aware routing[10], [11] and/or distributed medium access control[12]. In this paper we address the following questions:

- 1) How to optimally arrange distributed compression subject to aggregation costs to a *single* sink?
- 2) How to optimally arrange hierarchical architectures for aggregation/compression in large-scale sensor networks including *multiple* sinks?

Our aim is to minimize *overall* aggregation costs, i.e., energy expenditures, associated with gathering sensor information. In practice the spatial distribution of such expenditures may be critical. Indeed if nodes do not have renewable sources of energy, their batteries may become quickly depleted. Thus, as in [13], it also makes sense to address the above questions with a view at maximizing the network's lifetime, rather than minimizing the overall energy expenditure. In the sequel we will argue that the two cost functions are not 'orthogonal' and solutions minimizing overall energy consumption can be used as a good starting point towards maximizing network lifetime.

There has been much related work in this area. In particular we will draw on a substantial body of work studying the scaling and possible implementation of distributed compression mechanisms for sensor networks, e.g., see [5], [6], [7]. Our main contribution to this literature is to explicitly introduce aggregation costs in the distributed compression problem. The work of [14] is also closely related to ours. They propose a particular organization, based on proximity, of sensors and cluster-heads leading to a Voronoi tessellation of the sensing field. Their goal is to optimize the size of clusters so that the overall energy consumption of the network is minimized. By contrast, in this paper we attempt to devise an optimal hierarchical organization of sensors, aggregation points/compressors, and sinks, to minimize the cost of gathering sensor data.

The organization and main contributions of this paper can be summarized as follows. In Section II we formulate and solve an optimal distributed compression problem subject to aggregation, e.g., energy, costs to a single sink. In Section III and Section IV we propose a model for a large scale sensor network including multiple sinks, compression/aggregation devices and sensors, and determine the hierarchical organization resulting in a minimum overall energy consumption.

Manuscript received July 15, 2003; revised February 01, 2004. This work was supported by National Science Foundation under Grant ECS-0225448.

S. Baek and G. de Veciana are with the Department of Electrical and Computer Engineering, University of Texas, Austin, TX 78712 USA (e-mail: sbaek@ece.utexas.edu; gustavo@ece.utexas.edu).

X.Su is with Dept. of High Energy Physics, California Institute of Technology, Pasadena, CA 91125 USA (e-mail: xsu@hep.caltech.edu)

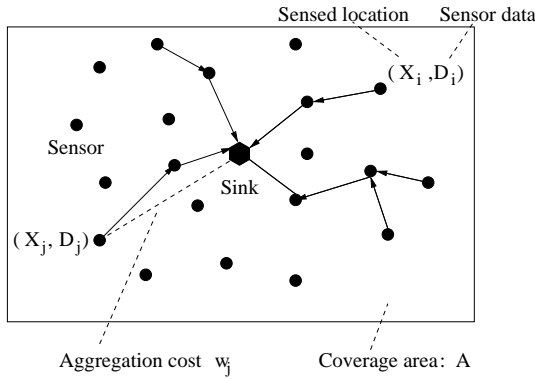


Fig. 1. An illustration of the sensor reporting.

Leveraging previous work on stochastic geometry, we estimate and compare the costs associated with various organizational structures including those based on simply relaying traffic to the closest compressor or sink. These provide insights on where these structures are likely to be effective. In Section V we provide numerical and simulation results for the energy savings obtained by using the optimal hierarchical scheme and discuss some design issues associated with the large-scale sensor networks. We conclude the paper in Section VI.

## II. OPTIMAL DISTRIBUTED DATA COMPRESSION SUBJECT TO AGGREGATION COSTS TO A SINGLE SINK

We begin by examining how to best realize distributed data compression subject to aggregation costs to a single sink. As illustrated in Fig.1, we consider a set of sensors  $U = \{1, 2, \dots, n\}$  at locations  $\vec{x} = (x_i, i \in U)$  within a coverage area  $A \subset \mathbb{R}^2$  and sink placed at the origin. We model information obtained by these sensors as a random vector  $\vec{D} = (D_i, i \in U)$ , and suppose the sink coordinates and aggregates the information from the sensors. Since sensors are located at different positions, they may incur different aggregation costs in forwarding their data to the sink. We use a vector  $\vec{w} = (w_i, i \in U)$  to model these costs, where  $w_i$  is the cost of moving a unit of data from sensor  $i$  to the sink. For example, the aggregation costs can be modelled by  $w_i = |x_i|^\gamma$ , i.e., a function of the distance from sensor  $i$  to the sink. Thus for  $\gamma = 1$  the cost for sensor  $i$  is proportional to its distance (a rough estimate for the number of hops) to the sink. Also if the sensor communicates directly with the sink then for  $\gamma \in (2, 4)$  this cost might capture the increased transmit power levels required to overcome path loss on a wireless link.

The information collected by the sensors is likely to be correlated and thus it is possible to jointly compress the data they generate. Let  $\vec{r} = (r_i, i \in U)$  denote the number of bits per reading each sensor would generate. Then by Slepian-Wolf's Theorem, the sum of the rates for any subset of sensors  $S \subset U$  is lower bounded by conditional entropy,  $H(D_S|D_{S^c})$ , where  $D_S = (D_i, i \in S)$  and  $D_{S^c} = (D_i, i \in U \setminus S)$  [15]. Our objective is to jointly compress the sensed data while minimizing the overall aggregation cost. We can formally state the problem as follows.

*Problem 1: For a set of devices  $U$  sensing an information vector  $\vec{D}$ , and an associated aggregation cost vector  $\vec{w}$  deter-*

*mine the rate vector  $\vec{r}^*$  that minimizes the overall aggregation cost subject to joint data compression constraints, i.e.,*

$$\min_{\vec{r}} \left\{ \sum_{i=1}^n w_i \cdot r_i \mid \sum_{i \in S} r_i \geq H(D_S|D_{S^c}), \forall S \subseteq U \right\}.$$

Note that the feasible rate region is defined by  $2^n - 1$  inequality constraints. Fortunately, one can show that the feasible rate region has a *contra-polymatroid* structure leading to the following greedy solution.

*Theorem 1: Suppose  $w_1 \geq w_2 \geq \dots \geq w_n$ , then the optimal solution to Problem 1 is given by*

$$\begin{aligned} r_1^* &= H(D_1|D_2, D_3, \dots, D_n), \\ r_2^* &= H(D_2|D_3, D_4, \dots, D_n), \dots, r_n^* = H(D_n). \end{aligned}$$

A *contra-polymatroid* is a polyhedron defined as follows:

$$\mathcal{B}(f) = \left\{ \vec{x} \mid \vec{x} \in \mathbb{R}_+^n, \sum_{i \in S} x_i \geq f(S), \forall S \subseteq U \right\},$$

where  $f: 2^U \rightarrow \mathbb{R}_+$  is called the rank function satisfying

- (1)  $f(\emptyset) = 0$ ;
- (2)  $f(S) \leq f(T) \quad \forall S \subset T$ ; (monotonicity)
- (3)  $f(S) + f(T) \leq f(S \cup T) + f(S \cap T)$  (super-modularity).

For such constraint sets the following result has been proved, see e.g., [16].

*Lemma 1: A solution  $\vec{r}^*$  to the following optimization problem where  $\vec{w} \in \mathbb{R}_+^n$ ,*

$$\min_{\vec{r}} \left\{ \sum_{i=1}^n w_i \cdot r_i \mid \vec{r} \in \mathcal{B}(f) \right\},$$

*is given by  $r_{\pi(1)}^* = f(\{\pi(1)\})$  and  $r_{\pi(i)}^* = f(\{\pi(1), \pi(2), \dots, \pi(i)\}) - f(\{\pi(1), \pi(2), \dots, \pi(i-1)\})$ ,  $i = 2, 3, \dots, n$ , where  $\pi$  denotes a permutation of  $(1, 2, \dots, n)$  such that  $w_{\pi(1)} \geq w_{\pi(2)} \geq \dots \geq w_{\pi(n)}$ .*

The key step to proving Theorem 1 is showing that the conditional entropy function  $f(S) = H(D_S|D_{S^c})$  is a rank function so the constraints in Problem 1 define a *contra-polymatroid*. The result then follows immediately by Lemma 1. The derivation of this result follows by a fairly straightforward argument—we refer the readers to [17] for a detailed proof.

Theorem 1 implies that a data compression strategy that minimizes the aggregation cost can be obtained by a simple ordering of the sources based on their aggregation cost. Surprisingly, the optimal solution is independent of the underlying correlation structure of the data sources or absolute values for these costs. Also note that Slepian-Wolf's result generalizes to ergodic data sources[15], so Thm.1 easily extends to data sources which are spatio-temporally correlated, by replacing conditional entropies with conditional entropy rates.

Let us briefly evaluate the performance gains that might be achieved under optimal compression. Suppose we deploy 10 sensors at a set of randomly selected locations  $\vec{x}$  in a square area  $A = [-1, 1]^2$ , with a sink located at the origin. We shall model the sensed data as a stationary Gaussian field with zero mean and a spatial covariance  $R(x_i, x_j) = 10e^{-\theta|x_i - x_j|}$

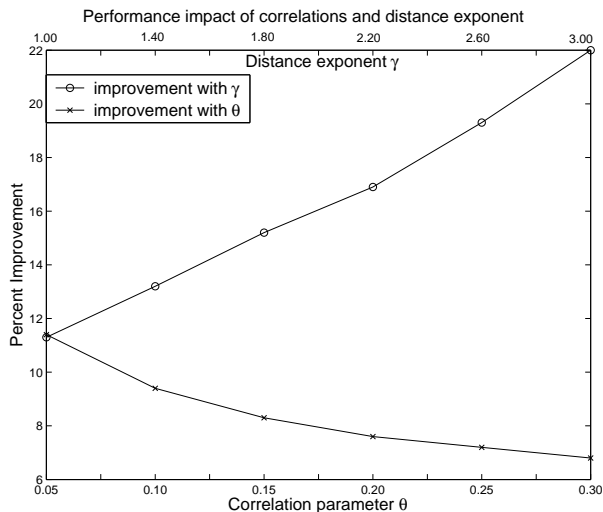


Fig. 2. The figure exhibits the percent energy savings as a function of the correlation  $\theta$  among data sources, and percent energy savings as a function of  $\gamma$  when  $\theta = 0.1$ .

where  $\theta$  models the rate of decay in the correlation between sensed data  $D_i$  and  $D_j$ , as a function of the distance between sensor locations  $x_i$  and  $x_j$ . This implies a correlation structure for the sensor data  $\vec{D}$  that is dependent on the locations of the sensors. To estimate the rates generated by sensors we suppose they are equipped with a 4-bit A/D converter, and thus approximately  $H(D_i|D_{i+1}, D_{i+2}, \dots, D_n) = 4 + h(D_i|D_{i+1}, D_{i+2}, \dots, D_n)$  where  $h(\cdot|\cdot)$  is the conditional differential entropy of the Gaussian source.

We will compare the *percentage reduction* in the aggregation cost for our optimal scheme versus a baseline which achieves optimal compression but allocates bits based on a random ordering of sensors. Fig.2 shows the case where the weight for sensor  $i$  depends on its distance to the sink  $|x_i|$ . As expected, the benefits decrease as correlation across sensed data decreases. This is intuitive since eventually with no correlation, i.e., independence, there will be no benefit for optimal distributed compression. Fig.2 also shows the performance improvement when the weights are given by  $|x_i|^\gamma$  and  $\gamma \in [1, 3]$ . As might be expected, larger exponents  $\gamma$  lead to higher differences in the aggregation costs among sensors and thus higher the performance gains.

### III. OPTIMAL HIERARCHICAL STRUCTURES FOR COMPRESSION AND AGGREGATION TO MULTIPLE SINKS

The scheme discussed in the previous section will not scale as the number sensors grows. Specifically, if the overall data produced increases, it will eventually overwhelm the sink. A more reasonable scenario for a large scale sensor network, is one with a three level hierarchy, including: at the top multiple sinks, e.g. wired nodes; at an intermediate level, a class (sensor) nodes that serve as intermediate traffic aggregators and/or in-network compressors; and at the bottom a collection of spatially distributed sensors. Each intermediate node collects data from an associated set of sensors and forwards it to a sink. Thus such nodes may minimize the energy costs to gather data

using optimal distributed compression. In addition, in practice, such nodes might play a key role in further compressing data sent to the sinks through aggregation/header compression. We will focus on this second role in this section. Our objective is to study how to best arrange such hierarchies so as to minimize the overall energy costs. Due to the complexity of and spatial character of problem, we will use the methodology proposed in [18]. The idea is to use crude stochastic geometric models to capture the salient features of the system.

#### A. Hierarchical organization and energy models

We shall assume the sensor, compressor and sink locations are determined by homogeneous Poisson point processes  $\Pi_0, \Pi_1$  and  $\Pi_2$ , with intensities  $\lambda_0, \lambda_1$  and  $\lambda_2$  respectively. We also assume that sinks may serve as compressors, i.e., it may gather data directly from sensors if necessary. We use the location of a node as an index for the device itself. Thus  $x_i \in \Pi_0$  denotes the sensor at location  $x_i$ ,  $c_j \in \Pi_1$  a compressor at location  $c_j$  and  $s_k \in \Pi_2$  a sink at location  $s_k$ . Our model for a hierarchical organization of these nodes is based on two mappings. Let  $c : \Pi_0 \rightarrow \Pi_1 \cup \Pi_2$  be a mapping associating a sensor  $x_i$  with a compressor or sink  $c(x_i)$  and  $s : \Pi_1 \cup \Pi_2 \rightarrow \Pi_2$  be a mapping associating a compressor or sink  $y_i$  with a sink  $s(y_i)$ . If  $y_i$  is a sink it is associated with itself. In the sequel we will consider various organizational structures, i.e., various choices for the functions  $c$  and  $s$ .

Next we propose a rough model for the energy costs associated with aggregation. We will assume that sensors and compressors communicate over an ad hoc network of wireless links, so packets will be relayed over multiple hops toward their destination. Thus the overall aggregation cost in the system will be proportional to the number of packets that need to be relayed. The energy cost per packet will depend on the overheads to access the channel for transmission, and energy expended during transmission. The latter would in turn depend on the size of the packet and the distances involved. We shall assume that packets have roughly the same size. Also assuming the devices are distributed in a fairly homogenous manner it is reasonable to assume the energy cost per packet relayed in the network is roughly the same, and the number of hops required to travel between two locations is proportional to distance between them. Suppose then that sensor  $x_i$  generates  $d_{x_i}$  packets/sec to its compressor  $c_j = c(x_i)$ . The total traffic received by  $c_j$  would be

$$\sum_{\{x_i:c(x_i)=c_j\}} d_{x_i} \text{ packets/sec,}$$

and the total energy expended in the network to carry this traffic is roughly proportional to

$$\sum_{\{x_i:c(x_i)=c_j\}} d_{x_i} |x_i - c_j| \text{ packets-hops/sec.}$$

Recall that each compressor  $c_j$  aggregates incoming data and possibly further compresses it prior to forwarding it to the sink  $s(c_j)$ . Aggregation may make use of context dependent information to compress the data it forwards to the sink. For example, if sensors are fixed a sink will eventually know which

sensors are associated with a compressor, and it in turn need not forward location information for each data sample. We will assume that  $c_j$  forwards packets at a rate

$$\alpha_{c_j} \sum_{\{x_i: c(x_i)=c_j\}} d_{x_i} \text{ packets/sec}$$

to its associated sink  $s_k = s(c_j)$ , incurring an additional energy cost which is proportional to

$$\alpha_{c_j} \sum_{\{x_i: c(x_i)=c_j\}} d_{x_i} |c_j - s_k| \text{ packets-hops/sec}$$

The parameter  $\alpha_{c_j} \in [0, 1]$  captures the gross benefit of aggregation and compression resulting from  $c_j$ 's processing and forwarding of sensor data.

Our goal is to capture the salient features of such a hierarchy, so it makes sense to assume that each sensor nodes offer roughly the same traffic load, i.e., without loss of generality we let  $d_{x_i} = 1$  for all  $x_i \in \Pi_0$ , possibly representing an average load per sensor to its aggregation point after distributed compression. In addition, assuming that the set of sensors associated with a given compressor node is sufficiently large, and variations in  $\alpha_{c_j}$  will not be significant, we will assume that all compressor/aggregating nodes are equally effective, i.e.,  $\alpha_{c_j} = \alpha$  for all compressor nodes  $c_j \in \Pi_1$ . With these two assumptions we obtain an additive energy cost model, with a cost per sensor,  $e(x_i)$ , which is proportional to

$$e(x_i) = |x_i - c(x_i)| + \alpha |c(x_i) - s(c(x_i))|. \quad (1)$$

Note that, depending on the relative distances, in practice it may indeed be more efficient for nodes to communicate directly with aggregation points or sinks. Our model can be generalized to capture the direct communication among sensors, compressors and sinks. In this case, an energy cost of  $e^\gamma(x_i) = |x_i - c(x_i)|^\gamma + \alpha |c(x_i) - s(c(x_i))|^\gamma$ , could be used where  $\gamma$  represents the path loss exponent of the environment. We refer the reader to [17] for results and analysis for of this generalization.

### B. Optimal sensor hierarchies and spatial tessellations

Our objective is, given our three point processes and a fixed gross compression ratio  $\alpha$ , to determine the hierarchical organization, i.e. the functions  $c(\cdot)$  and  $s(\cdot)$ , that minimize the overall energy cost in this network. Prior to considering optimal hierarchies let us consider a natural choice. It should be clear that from the perspective of minimizing our cost metric, each compressor should relay packets to the closest sink. Similarly one might think it reasonable to assume each sensor sends its traffic to the closest compressor or, if it is closer, directly to a sink. As discussed in [18] such a hierarchical organization is induced by the *Voronoi tessellation* generated by the locations of the sinks and compressors. A Voronoi tessellation with respect to a set of points  $\Pi$  can be defined as follows.

*Definition 1:* The Voronoi tessellation associated with  $\Pi$ , denoted  $\mathcal{V}(\Pi)$  is a collection of cells  $V_{y_i}(\Pi)$  for  $y_i \in \Pi$  such that

$$V_{y_i}(\Pi) = \{z \in \mathbb{R}^2 \mid |y_i - z| < |y_j - z|, \forall y_j \in \Pi\},$$

i.e., all points in the plane which are closer to  $y_i$  than to any other point in  $\Pi$ .

Let  $\mathcal{V}(\Pi_1 \cup \Pi_2)$  denote the Voronoi tessellation induced by the union of compressors and sinks  $\Pi_1 \cup \Pi_2$  and let us define a hierarchical organization as follows. Let  $c_V : \Pi_0 \rightarrow \Pi_1 \cup \Pi_2$  be such that for each  $x_i \in \Pi_0$ ,

$$c_V(x_i) = y_j, \text{ iff } x_i \in \bar{V}_{y_j}(\Pi_1 \cup \Pi_2).$$

Here  $\bar{V}_{y_j}$  denotes the closure of the set  $V_{y_j}$ . Mathematically the probability that a sensor will fall on the boundary of sets  $V_{y_i}$  vanishes. However, if a sensor is observed to be located on the boundary due to the finite granularity in the positions of nodes, such ties must be broken arbitrarily. Let  $\mathcal{V}(\Pi_2)$  denote the Voronoi tessellation induced by the sinks  $\Pi_2$ , and define a mapping  $s_V : \Pi_1 \cup \Pi_2 \rightarrow \Pi_2$ , such that for each  $y_j \in \Pi_1 \cup \Pi_2$ ,

$$s_V(y_j) = s_k, \text{ iff } y_j \in \bar{V}_{s_k}(\Pi_2).$$

The pair  $c_V$  and  $s_V$  capture a simple hierarchical organization for sensors, compressors and sinks based purely on proximity rather than achieving our goal, i.e., minimizing the energy cost of aggregation. In particular a sensor may connect to the closest compressor even if the energy cost would be reduced if it connected to another compressor which is closer to its eventual sink and thus results in an energy savings. Below we define the tessellation  $\mathcal{T}^\alpha(\Pi_1, \Pi_2)$  that induces an optimal organization for a given overall compression ratio  $\alpha$ .

*Definition 2:* The tessellation  $\mathcal{T}^\alpha(\Pi_1, \Pi_2)$  associated with two sets of points  $\Pi_1$  and  $\Pi_2$  and parameter  $\alpha$ , consists of a collection of cells  $T_{y_i}^\alpha(\Pi_1, \Pi_2)$  for  $y_i \in \Pi_1 \cup \Pi_2$  such that

$$\begin{aligned} T_{y_j}^\alpha(\Pi_1, \Pi_2) &= \{z \in \mathbb{R}^2 \mid |z - y_j| + \alpha |y_j - s_V(y_j)| \\ &< |z - y_l| + \alpha |y_l - s_V(y_l)|, \forall y_l \in \Pi_1 \cup \Pi_2\}. \end{aligned}$$

We refer to  $T_{y_j}^\alpha(\Pi_1, \Pi_2)$  as the cell with nucleus  $y_j$ .

The next theorem shows that this tessellation induces an organization which minimizes the energy cost for each sensor.

*Theorem 2:* Under the energy cost model (1), the optimal assignment of compressors  $c_j \in \Pi_1$  to sinks is  $s^*(c_j) = s_V(c_j) \in \Pi_2$ , i.e., the closest sink to  $c_j$ . The optimal assignment of sensors  $x_i \in \Pi_0$  to compressors or sinks is  $c^*(x_i)$  such that  $x_i \in T_{c^*(x_i)}^\alpha(\Pi_1, \Pi_2) \in \mathcal{T}^\alpha(\Pi_1, \Pi_2)$ , i.e., the compressor/sink which is the nucleus for the cell in  $\mathcal{T}^\alpha(\Pi_1, \Pi_2)$  containing  $x_i$ .

*Proof:* To prove that this hierarchy has minimal cost consider a sensor  $x_i \in T_{y_j}^\alpha(\Pi_1, \Pi_2)$ , so  $c^*(x_i) = y_j$ . Suppose  $x_i$  is assigned to some other compressor or sink  $y_l$  and then some sink  $s_k$ . The energy cost for such an assignment would be

$$\begin{aligned} e(x_i) &= |x_i - y_l| + \alpha |y_l - s_k| \\ &\geq |x_i - y_l| + \alpha |y_l - s^*(y_l)| \\ &\geq |x_i - y_j| + \alpha |y_j - s^*(y_j)| \\ &= |x_i - c^*(x_i)| + \alpha |c^*(x_i) - s^*(c^*(x_i))|, \end{aligned}$$

where the first inequality follows from the definition of  $s^*$  and the second inequality is a natural consequence of the definition of the cells in the tessellation  $\mathcal{T}^\alpha(\Pi_1, \Pi_2)$ . ■

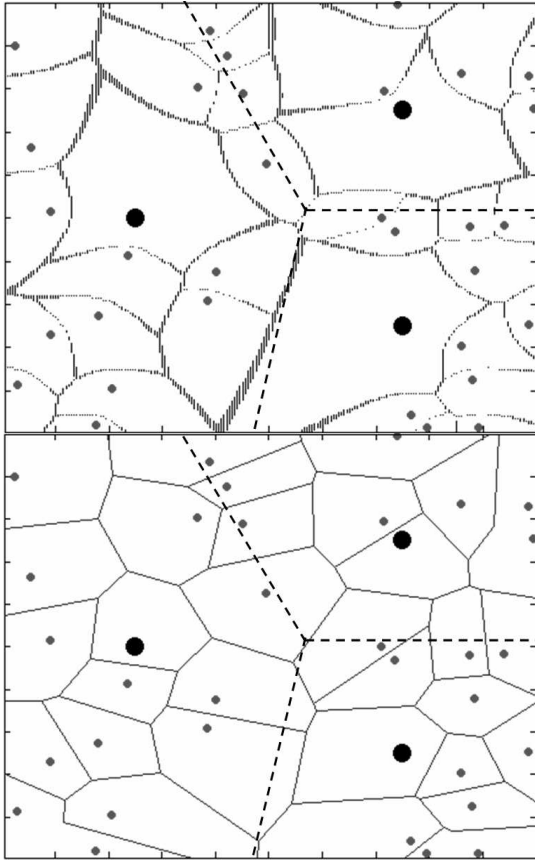


Fig. 3. The figure on the above shows the optimal Johnson-Mehl tessellation  $\mathcal{T}^\alpha(\Pi_1, \Pi_2)$  for compression ratio  $\alpha = 0.5$ , and the figure on the below shows the Voronoi tessellation  $\mathcal{V}(\Pi_1 \cup \Pi_2)$ . The small dots represent compressors in  $\Pi_1$  and big dots represent sinks in  $\Pi_2$ . We have superposed the Voronoi tessellation  $\mathcal{V}(\Pi_2)$  induced by the sinks and drawn the boundaries with dashed straight lines. The organization of sensors to compressors or sinks, and compressors to sinks should be clear based on its location with respect to the two superposed tessellations.

The tessellation  $\mathcal{T}^\alpha(\Pi_1, \Pi_2)$  introduced earlier is a particular case of a Johnson-Mehl tessellation [19]. Let us briefly consider how this comes about by contrasting it with a Voronoi tessellation. One can visualize the Voronoi tessellation as induced by a set of nuclei which grow grains isotropically at unit speed until they come into contact with another nuclei's grain. These contact points become the boundaries of the cells of the tessellation. A Johnson-Mehl tessellation adds an additional degree of freedom. It allows the nuclei to start growing their grains at different times. The tessellation  $\mathcal{T}^\alpha(\Pi_1, \Pi_2)$  results if all sinks  $s_k \in \Pi_2$  start growing grains together isotropically at unit speed. Subsequently each compressor  $c_j \in \Pi_1$  starts growing a grain at time proportional to  $\alpha|c_j - s^*(c_j)|$  which depends on its distance from the closest sink and  $\alpha$ . The resulting boundaries can be shown to be hyperbolic. In this tessellation, the cells associated with sinks will tend to be large relative to others, since they start growing earlier. For a formal proof of this argument, we refer the readers to [17]. Examples of a Voronoi tessellation and an energy optimal Johnson-Mehl tessellation for  $\alpha = 0.5$  are shown in Figure 3.

#### IV. ANALYTICAL RESULTS

In this section we will compare various organizational hierarchies in terms of their expected energy costs.

##### A. Hierarchy based on Voronoi tessellation

Let us first consider the hierarchy based on Voronoi tessellation whereby each sensor sends data to its closest compressor or sink node, and compressor nodes relay a fraction  $\alpha$  of the traffic they receive to the closest sink. Following [18] we can mathematically define an expected energy cost  $G_V$  under such an arrangement as

$$G_V = \mathbf{E}_2^0 \left[ \sum_{c_j \in \Pi_1 \cap V_0(\Pi_2)} \{\alpha|c_j| \mathcal{N}_{c_j} + \sum_{x_i \in \Pi_0 \cap V_{c_j}(\Pi_1 \cup \Pi_2)} |x_i - c_j|\} + \sum_{x_i \in \Pi_0 \cap V_0(\Pi_1 \cup \Pi_2)} |x_i| \right],$$

where we have used the following conventions:  $\mathbf{E}_k^0$  denotes the expectation with respect to the Palm distribution  $\mathbf{P}_k^0$  of the point process  $\Pi_k$ , see [20]; and  $\mathcal{N}_{c_j}$  is the number of sensors, i.e., points in  $\Pi_0$  which are associated with compressor  $c_j$ , that is in the cell  $V_{c_j}(\Pi_1)$ .

Thus  $G_V$  corresponds to the energy cost associated with a typical sink cell  $V_0(\Pi_2)$ . The first summand consists of two terms: the first corresponds to the energy costs in forwarding compressed data from compressors to sinks, while the second corresponds to the energy cost of forwarding data from sensors to compressors. The second summand in the above expression corresponds to energy costs of sensors that directly send data to the sink. One can use the results in [18] to evaluate the expected cost  $G_V$  giving<sup>1</sup>:

$$G_V = \frac{\lambda_0}{2\lambda_2\sqrt{\lambda_1 + \lambda_2}} + \frac{\alpha\lambda_0}{2\lambda_2^{3/2}} - \frac{\alpha\lambda_0}{2(\frac{\lambda_1}{4} + \lambda_2)^{3/2}} + \int_0^{\frac{\pi}{2}} \frac{12\pi^{3/2}\lambda_0\lambda_1\alpha(\pi - \gamma)\sin\gamma\cos^2\gamma d\gamma}{(\pi(\gamma_1 + 2\gamma_2) + \gamma_2\sin(2\gamma) + 2\gamma_2(\pi - \gamma)\cos(2\gamma))^{5/2}}.$$

Note that the same results can be applied to evaluate the cost of an aggregation hierarchy with no compressor nodes at all, i.e., sensors merely send uncompressed data to the closest sink. Applying the results in [18] one finds the energy cost  $G_S$  for a typical sink cell is given by

$$G_S = \frac{\lambda_0}{\lambda_2} \cdot \frac{1}{2\sqrt{\lambda_2}}.$$

This can be interpreted as follows: the area of a typical sink cell is  $1/\lambda_2$  and so the average number of sensors in that cell is expected to be  $\lambda_0/\lambda_2$ . Moreover one can show that  $1/2\sqrt{\lambda_2}$  corresponds to the mean distance from a sensor to its closest sink. Thus the average cost should be the product of these two terms.

As will be seen in the sequel, it may be the case that  $G_V > G_S$  for some  $\alpha$ . Specifically when  $\alpha$  is close to 1 it may be more expensive to route to a compressor which is close by

<sup>1</sup>We have corrected an error in the derivations of [18] to get this expression for  $G_V$ .

rather than go directly to a sink. Thus a greedy aggregation strategy based purely on proximity may be worse than not using one at all.

### B. Optimal hierarchy based on Johnson-Mehl tessellation

Next we evaluate the energy cost  $G_T$  for the optimal hierarchy introduced in the previous section. It is given by

$$G_T = \mathbf{E}_2^0 \left[ \sum_{c_j \in \Pi_1 \cap V_0(\Pi_2)} \{\alpha |c_j| \mathcal{N}_{c_j}^\alpha\} + \sum_{x_i \in \Pi_0 \cap T_{c_j}^\alpha(\Pi_1, \Pi_2)} |x_i - c_j| \} + \sum_{x_i \in \Pi_0 \cap T_0^\alpha(\Pi_1, \Pi_2)} |x_i| \right],$$

where  $\mathcal{N}_{c_j}^\alpha$  denotes the number of sensors, i.e., points in  $\Pi_0$ , which are associated with compressor  $c_j$ , and are within the cell  $T_{c_j}^\alpha(\Pi_1, \Pi_2)$ . Here  $G_T$  corresponds to the energy cost associated with a typical cell  $V_0(\Pi_2)$  in the higher level Voronoi tessellation induced by sinks.

As an intermediate step toward estimating the energy cost, let us consider the area of a typical *sink* cell in the Johnson-Mehl tessellation  $\mathcal{T}^\alpha(\Pi_1, \Pi_2)$ . Recall that cells associated with sinks will tend to be larger than others since they start growing earlier, see Fig. 3. The following theorem, proved in the appendix, gives an explicit formula for the area of a typical sink.

*Theorem 3: Consider two point processes  $\Pi_1$  (compressors) and  $\Pi_2$  (sinks) with densities  $\lambda_1$  and  $\lambda_2$  respectively. The expected area of a typical sink cell in the tessellation  $\mathcal{T}^\alpha(\Pi_1, \Pi_2)$  is given by*

$$\mathbf{E}_2^0[ |T_0^\alpha(\Pi_1, \Pi_2)| ] = \frac{\pi}{f(\alpha)\lambda_1 + \pi\lambda_2}$$

where  $|T|$  denotes the area of a set  $T$  and  $f(\alpha)$  is defined as follows. We let  $O_\beta^\alpha(x_1, x_2)$  denote the Cartesian oval with foci at  $x_1, x_2 \in \mathbb{R}^2$ , given by

$$O_\beta^\alpha(x_1, x_2) = \{y \in \mathbb{R}^2 \mid \alpha|y - x_1| + |y - x_2| \leq \beta\},$$

and let  $f(\alpha)$  be the area of  $O_1^\alpha(0, x)$  when  $|x| = 1$ . Then  $f(\alpha)$  is given by

$$f(\alpha) = \frac{4}{(1 - \alpha^2)^2} \left\{ \left( \frac{1}{2} + \alpha^2 \right) \left( \frac{\pi}{2} - \sin^{-1}(\alpha) \right) + \frac{\sin(2 \sin^{-1}(\alpha))}{4} - 2\alpha \cos(\sin^{-1}(\alpha)) \right\}. \quad (2)$$

Note that the Cartesian oval  $O_\beta^\alpha(0, x)$  can be interpreted as the set of possible compressor locations with respect to a sink at the origin and a sensor at  $x$  that would result in a lower energy cost than  $\beta$  if used as a relay point, see e.g., Fig. 4. One can show that  $f(\alpha)$  is a monotonically decreasing function in  $\alpha \in [0, 1]$  where  $f(0) = \pi$ ,  $f(1) = 0$ . Thus when  $\alpha = 1$ , i.e., there is no compression gain, the area of a typical sink cell becomes  $\frac{1}{\lambda_2}$ . This is the average area of a typical cell for a Voronoi tessellation with density  $\lambda_2$ , thus connecting to the nearest sink must be optimal. At the other extreme if  $\alpha = 0$ , and thus the data which will be relayed from compressors to sinks is negligible, then the

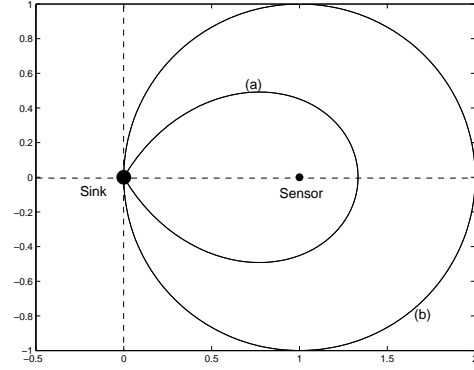


Fig. 4. The figure shows the ball  $B_{|x|}(x)$  and Cartesian oval  $O_{|x|}^\alpha(0, x)$  associated with a potential (sensor) location at  $x = (1, 0)$  and sink at the origin. In order for the sensor at  $(1, 0)$  to be associated with the sink at the origin, there must be no compressors in the Cartesian oval (a) and no other sinks in region in the ball (b).

optimal sink cell's area becomes  $\frac{1}{\lambda_1 + \lambda_2}$  which is the average area of a typical cell of a Voronoi tessellation with density  $\lambda_1 + \lambda_2$ . Clearly connecting to the nearest sink or compressor is the optimal policy when  $\alpha = 0$ . Thus at the two extremes our Johnson-Mehl tessellation reduces to one of two Voronoi cases. However when  $\alpha \in (0, 1)$ , we can expect to get cells with hyperbolic faces associated with an optimal hierarchy.

Unfortunately since  $\mathcal{T}^\alpha(\Pi_1, \Pi_2)$  depends in an inhomogeneous manner on the two point processes, estimating the energy cost is quite challenging. As such, we derive an upper bound  $G_U$  for the optimal energy cost  $G_T$ . To do so, consider yet another alternative hierarchical organization for the sensor network. In this hierarchy the data associated with a sensor  $x_i$  is necessarily routed to its closest sink  $s_V(x_i)$ . This data may however be routed either directly to the sink or via a compressor depending on which option minimizes the contribution to the energy cost of the system. Letting  $c_U(x_i)$  denote the energy optimal routing choice for sensor  $x_i$ , i.e., to a sink or compressor, the associated energy cost is given by

$$e(x_i) = |x_i - c_U(x_i)| + \alpha |c_U(x_i) - s_V(x_i)|. \quad (3)$$

Although this hierarchy is suboptimal it is quite close to the optimal one. In particular sensors which are not close to the boundary of the Voronoi cells associated with the point process of sinks, are likely to contribute the same energy costs as in the optimal hierarchy.

We can define the energy cost for a typical cell  $G_U$  for this new hierarchical arrangement as

$$G_U = \mathbf{E}_2^0 \left[ \sum_{x_j \in \Pi_0 \cap V_0(\Pi_2)} e_U(x_j, \Pi_1, \Pi_2) \right].$$

Under this new organization only sensors within  $V_0(\Pi_2)$  will be associated with the origin, thus one can directly express the energy cost of a typical cell as a sum over sensors in this cell. However the cost per sensor depends on the locations of compressor and sink nodes  $\Pi_1$  and  $\Pi_2$  – we introduce  $e_U$  to make this explicit.

The following theorem, proved in the appendix, gives an

explicit upper bound for the energy cost achieved by the optimal aggregation scheme.

*Theorem 4:* The energy  $G_U$  is an upper bound for the optimal cost  $G_T$  and is given by

$$G_U = \frac{\lambda_0 \pi \Gamma(3/2)}{(\lambda_2 \pi + \lambda_1 f(\alpha))^{3/2}} + \lambda_0 \lambda_1 \int_{\mathbb{R}^2} e^{-\lambda_2 \pi |x|^2} \int_{O_{|x|}^\alpha(x,0)} (|y| + \alpha |y-x|) e^{-\lambda_1 |O_{|y|+\alpha|y-x|}^\alpha(x,0)|} dy dx \quad (4)$$

where  $f(\alpha)$  and  $O_\beta^\alpha(x,0)$  are as defined in Theorem 3. Furthermore,  $G_U$  is upper bounded by  $\hat{G}_U$  given by

$$\hat{G}_U = \frac{\lambda_0}{\lambda_2} \left\{ \frac{\alpha}{2\sqrt{\lambda_2}} + \frac{1-\alpha}{2\sqrt{\lambda_2 + \lambda_{eff}}} \right\}, \quad (5)$$

where  $\lambda_{eff} = \frac{f(\alpha)}{\pi} \lambda_1$ .

By Neveu exchange formula [18], one can interpret the expression in the bracket of (5) as an average cost incurred by a typical sensor. Observe that for perfect compression, i.e.,  $\alpha = 0$ , we have that  $f(0) = \pi$  so this term reduces to  $\frac{1}{2\sqrt{\lambda_1 + \lambda_2}}$  which corresponds to the typical cost of a sensor connecting to the closest sink or compression. By contrast for the case with no compression at all, i.e.,  $\alpha = 1$ , this term reduces to  $\frac{1}{2\sqrt{\lambda_2}}$ , the cost for a typical sensor to connect to its closest sink. This clarifies the dependence of the energy cost on the compression ratio  $\alpha$  as we move among these two extremes. Finally we note that through simulation we found that  $\hat{G}_U$  gave an excellent approximation for  $G_U$  –the error was less than 1.4%, see [17] for more details.

## V. NUMERICAL AND SIMULATION RESULTS

We numerically evaluated the energy costs for the various schemes discussed in the previous section. Namely  $G_S$  corresponds to the case where sensors send directly to the sink without going through compressors,  $G_V$  the case where sensors send to a compressor or a sink, whichever is closer, and then compressors send to sinks, and  $G_U$  is our upper bound for the optimal strategy. As seen at the top in Fig.5, the optimal hierarchy (upper bound  $G_U$ ) can outperform those based on Voronoi tessellation, i.e.  $G_S$  and  $G_V$ . With a moderate compression ratio, e.g.  $0.5 \leq \alpha \leq 0.9$ , the percent energy savings of the optimal scheme versus the Voronoi scheme  $G_V$  is 6%-26%. Also note that our closed-form approximation  $\hat{G}_U$  is a tight upper bound of  $G_U$ : the error is within 1.4% of  $G_U$ . We have verified by simulation that the percent energy savings is in fact roughly 8%-28%. However it is very close to the simulated cost, where the error is within 2%. Details of the simulation methodology and results can be found in [17].

Observe that as  $\alpha$  increases the second scheme  $G_V$  is inferior to  $G_S$ ; this is intuitive since  $G_V$  is constrained to connect to the closest compressor if there are no sinks which are closer, thus requiring a “detour” even when no compression gain is available. Note however, our optimal scheme always performs better than the other two. In particular as soon as the aggregation/compression nodes provide a reasonable compression ratio it is able to significantly reduce the overall energy cost of the network. The plot at the bottom of Fig.5

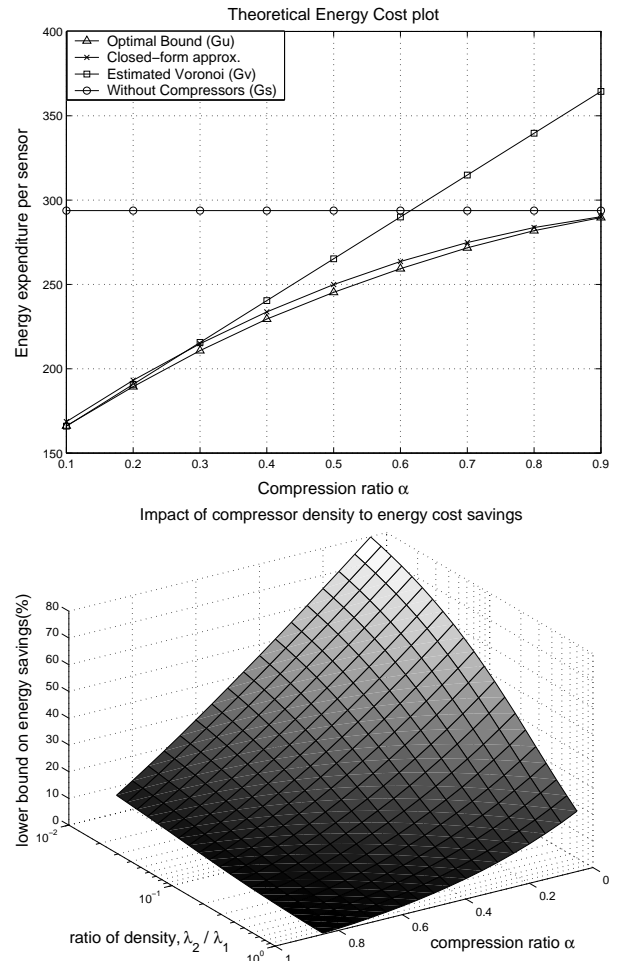


Fig. 5. The graph at the top exhibits a numerical evaluation of  $G_S$ ,  $G_V$  and our upper bounds  $G_U$  and  $\hat{G}_U$  for the optimal hierarchy, as a function of  $\alpha$  with  $\lambda_0 = 1$ ,  $\lambda_1 = 0.0473$  and  $\lambda_2 = 0.0142$ . The graph at the bottom exhibits energy savings, comparing  $G_U$  to  $G_S$ , as a function of  $\alpha$  and  $\lambda_2/\lambda_1$ .

shows the percent energy savings for  $G_U$  versus  $G_S$ , as a function of  $\alpha$  and  $\lambda_2/\lambda_1$ . Note the sensitivity of the savings for the optimal scheme increases with the density of compressor nodes. The intuition here is that for high densities of compressors when  $\alpha$  changes, sensors will have more choices for candidate compressor nodes leading to higher savings.

In Fig.6, we show the spatial distribution of the energy contribution  $e(x)$  for possible sensor locations under the optimal hierarchical arrangement. Bright regions indicate areas where sensors incur a higher energy cost to the overall system, and dark regions are the areas with less cost. We observe that the energy contours vary in a smooth way. The reason is that the optimal tessellation is constructed for each sensor to achieve the minimal cost at its location, and the cost function is a smooth function of the sensor positions in  $\mathbb{R}^2$ . By contrast, a similar plot for the organization based on the Voronoi scheme (omitted here) shows discontinuities at the cell boundaries since sensors associate with closest compressors (or sinks) rather than the minimum energy choices.

In practice it is of interest to examine the *traffic congestion*

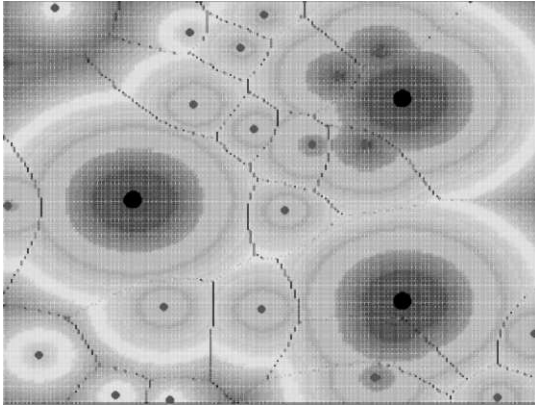


Fig. 6. Energy contribution plot for the optimal scheme when  $\alpha = 0.5$ . The big and small dots represent sinks and compressors respectively.

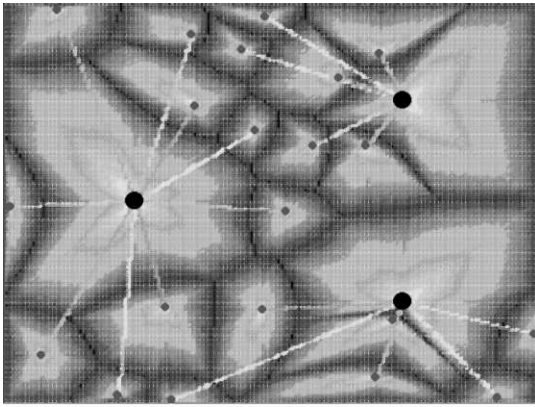


Fig. 7. Energy consumption plot for the optimal scheme when  $\alpha = 0.5$ . The big and small dots represent sinks and compressors respectively.

under various aggregation schemes. In particular this will reflect uneven energy burden throughout the network. Fig.7 exhibits a rough estimate for overall energy burden per unit transmission time assuming the sensor density is quite high and traffic is routed along straight lines. The figure clearly shows the high energy burden around the aggregation points, i.e. compressors or sinks, and imbalance of such burdens. As mentioned in the introduction, such imbalances may quickly exhaust the energy reserves of some nodes causing the network to eventually become disconnected or not provide the appropriate coverage. One way to alleviate this problem is to periodically update the positions of the compressors, e.g. [14], yet this will not resolve imbalances around sink nodes. Another way to alleviate this problem is to allow more flexible congestion or energy sensitive routing towards aggregation nodes [21]. We mainly discussed the optimal rate allocation and the optimal hierarchical organization schemes in this paper, and devising load balancing strategies combined with these schemes is part of our ongoing study.

## VI. CONCLUSION

Energy consumption will be a critical operational issue for wireless sensor networks. In this paper we first show how to

perform optimal distributed compression so as to minimize aggregation (energy) costs to a sink. We then propose a natural model for a large scale network including multiple sinks, compressor/aggregation nodes, and sensors. We study how to optimize this hierarchy so as to minimize the overall energy consumption— the optimal hierarchy is shown to be associated with a Johnson-Mehl tessellation induced by sinks and compressor nodes. Using stochastic geometry we were able to evaluate the average costs associated with such arrangements and compare them with natural proximity based organizations previously proposed. Our results suggest that there are significant energy savings, 8-28 %, to be achieved by optimizing compression and aggregation structures for sensor networks. However, minimizing the overall energy cost is but part of the picture. In practice, when sensors have non-replenishable energy sources, the objective will be to maximize the network lifetime, e.g., time until a sensor expires or a coverage constraint is violated. This is known to be an NP hard problem but has been partially tackled in [22], [13]. By combining our optimized hierarchy with more flexible energy sensitive routing schemes we believe that a fairly good solution to maximizing network lifetime can be achieved.

## APPENDIX

**Proof of Theorem 3:** We use techniques similar to those in [23]. The following sequence of equalities capture the key steps:

$$\mathbf{E}_2^0[|T_0^\alpha(\Pi_1, \Pi_2)|] = \mathbf{E}_2^0\left[\int_{\mathbb{R}^2} \mathbf{1}(x \in T_0^\alpha(\Pi_1, \Pi_2)) dx\right] \quad (6)$$

$$= \mathbf{E}_2^0\left[\int_{\mathbb{R}^2} \mathbf{1}(B_{|x|}(x) \cap \Pi_2 = \emptyset) \mathbf{1}(O_{|x|}^\alpha(0, x) \cap \Pi_1 = \emptyset) dx\right] \quad (7)$$

$$= \int_{\mathbb{R}^2} \mathbf{P}(B_{|x|}(x) \cap \Pi_2 = \emptyset) \mathbf{P}(O_{|x|}^\alpha(0, x) \cap \Pi_1 = \emptyset) dx \quad (8)$$

$$= \int_{\mathbb{R}^2} e^{-\pi\lambda_2|x|^2} e^{-\lambda_1 f(\alpha)|x|^2} dx = \frac{\pi}{f(\alpha)\lambda_1 + \pi\lambda_2}. \quad (9)$$

Eq.(6) is straightforward, since the area of a cell is the integral of the indicator function of the region. Eq.(7) follows by observing that location  $x$  will belong to the cell if two conditions are met. First there must be no additional sinks (in addition to the one at the origin) within ball  $B_{|x|}(x)$  of radius  $|x|$  centered at  $x$ . Second there must be no compressors within the Cartesian oval  $O_{|x|}^\alpha(0, x)$ , otherwise  $x$  would belong to the cell associated with that compressor rather than that corresponding to the sink at the origin. This second condition should be clear given our definition of the Cartesian oval and our cost function. These two requirements are shown in Fig. 4. Eq.(8) uses the fact that the Palm probability  $\mathbf{P}_2^0$  for a Poisson process  $\Pi_2$  is simply that of the stationary Poisson process with same intensity but with an additional point at the origin - Slivnyak's Theorem [24], [20]. In addition on Eq.(8) we use the fact that  $\Pi_1$  is independent of  $\Pi_2$ . Eq.(9) explicitly replaces the void probabilities in Eq.(8). Recall that for a Poisson process  $\Pi$  with intensity  $\lambda$  and set  $A$ , the probability of no points arising in  $A$  is simply  $\mathbb{P}(\Pi \cap A = \emptyset) = e^{-\lambda|A|}$ . Also the area of a Cartesian oval  $O_\beta^\alpha(0, x)$  in the special case



where  $\beta = |x|$  can be computed explicitly and is given by  $|O_{|x|}^\alpha(0, x)| = |x|^2 f(\alpha)$  where  $f(\alpha)$  is given by (2). The desired result Eq.(9) follows by direct integration.  $\blacksquare$

**Proof of Theorem 4:** By Neveu's exchange formula [18] we have that

$$\mathbf{E}_2^0 \left[ \sum_{x_j \in \Pi_0 \cap V_0(\Pi_2)} e_U(x_j, \Pi_1, \Pi_2) \right] = \frac{\lambda_0}{\lambda_2} \mathbf{E}_0^0 [e_U(0, \Pi_1, \Pi_2)]$$

where  $e_U(0, \Pi_1, \Pi_2)$  denotes the energy cost for typical sensor at the origin. This energy cost depends on whether our sensor chooses to send directly to the sink or through a compressor. We shall consider these cases separately. In order to send directly to a sink  $s_k \in \Pi_2$  it must be the case that  $s_k$  is the closest to the origin, i.e.,  $B_{|s_k|}(0) \cap \Pi_2 = \emptyset$  and there is no eligible compressor in the Cartesian oval  $O_{|s_k|}^\alpha(s_k, 0)$ , i.e.,  $O_{|s_k|}^\alpha(s_k, 0) \cap \Pi_1 = \emptyset$ . Indeed any compressor within this oval would offer a cost which is less than  $|s_k|$  and thus preclude direct connection to the sink  $s_k$ . We can compute the energy cost associated with this first case as follows:

$$\begin{aligned} & \mathbf{E}_0^0 \left[ \sum_{s_k \in \Pi_2} |s_k| \mathbf{1}(B_{|s_k|}(0) \cap \Pi_2 = \emptyset) \mathbf{1}(O_{|s_k|}^\alpha(s_k, 0) \cap \Pi_1 = \emptyset) \right] \\ &= \mathbf{E} \left[ \sum_{s_k \in \Pi_2} |s_k| \mathbf{1}(B_{|s_k|}(0) \cap \Pi_2 = \emptyset) \right. \\ & \quad \left. \times \mathbf{1}(O_{|s_k|}^\alpha(s_k, 0) \cap \Pi_1 = \emptyset) \right] \quad (10) \\ &= \lambda_2 \int_{\mathbb{R}^2} |x| \mathbf{P}_2^0(B_{|x|}(-x) \cap \Pi_2 = \emptyset) \\ & \quad \times \mathbf{P}(O_{|x|}^\alpha(-x, 0) \cap \Pi_1 = \emptyset) dx \quad (11) \\ &= \lambda_2 \int_{\mathbb{R}^2} |x| e^{-\lambda_2 \pi |x|^2 - \lambda_1 |O_{|x|}^\alpha(-x, 0)|} dx \quad (12) \\ &= 2\pi \lambda_2 \int_0^\infty r^2 e^{-(\lambda_2 \pi + \lambda_1 f(\alpha)) r^2} dr \quad (13) \\ &= \frac{\lambda_2 \pi \Gamma(3/2)}{(\lambda_2 \pi + \lambda_1 f(\alpha))^{3/2}}. \quad (14) \end{aligned}$$

Eq.(10) follows by the independence of  $\Pi_1$ ,  $\Pi_2$  and  $\Pi_0$ . Eq.(11) follows by the refined Campbell's Theorem [20]. In Eq.(12) we have explicitly replaced the void probabilities for the associated events. Eq.(13) corresponds to a change of variables, and finally Eq.(14) is an explicit computation of the integral.

The second case is more complex. In this case the sensor at the origin chooses to transmit to a compressor, say  $c_j$ , which then relays the data to the sink  $s_k$ . Note under the proposed hierarchy the sink must be the closest to the sensor, so  $B_{|s_k|}(0) \cap \Pi_2 = \emptyset$ . Note that in this case the energy cost will be  $|c_j| + \alpha|c_j - s_k|$  and the compressor leading to the minimum energy cost is selected. Thus it must be the case that  $O_{|c_j| + \alpha|c_j - s_k|}^\alpha(s_k, 0) \cap \Pi_1 = \emptyset$  i.e., there is no alternative compressor that can achieve a lower cost. Also one must ensure that  $|c_j| + \alpha|c_j - s_k| \leq |s_k|$  otherwise the sensor would minimize its energy cost by going directly to the sink. For simplicity let  $d(c_j, s_k) = |c_j| + \alpha|c_j - s_k|$ . We can compute cost incurred by sensors that choose to relay through

compressors as follows.

$$\begin{aligned} & \mathbf{E}_0^0 \left[ \sum_{s_k \in \Pi_2} \sum_{c_j \in \Pi_1} d(c_j, s_k) \mathbf{1}(B_{|s_k|}(0) \cap \Pi_2 = \emptyset) \right. \\ & \quad \left. \times \mathbf{1}(O_{d(c_j, s_k)}^\alpha(s_k, 0) \cap \Pi_1 = \emptyset) \mathbf{1}(d(c_j, s_k) \leq |s_k|) \right] \\ &= \mathbf{E} \left[ \sum_{s_k \in \Pi_2} \sum_{c_j \in \Pi_1} d(c_j, s_k) \mathbf{1}(B_{|s_k|}(0) \cap \Pi_2 = \emptyset) \right. \\ & \quad \left. \times \mathbf{1}(O_{d(c_j, s_k)}^\alpha(s_k, 0) \cap \Pi_1 = \emptyset) \mathbf{1}(d(c_j, s_k) \leq |s_k|) \right] \quad (15) \\ &= \lambda_2 \int_{\mathbb{R}^2} \mathbf{E}_2^0 \left[ \sum_{c_j \in \Pi_1} d(c_j, x) \mathbf{1}(B_{|x|}(-x) \cap \Pi_2 = \emptyset) \right. \\ & \quad \left. \times \mathbf{1}(O_{d(c_j, x)}^\alpha(x, 0) \cap \Pi_1 = \emptyset) \mathbf{1}(d(c_j, x) \leq |x|) \right] dx \quad (16) \\ &= \lambda_2 \int_{\mathbb{R}^2} e^{-\lambda_2 \pi |x|^2} \mathbf{E} \left[ \sum_{c_j \in \Pi_1} d(c_j, x) \right. \\ & \quad \left. \times \mathbf{1}(O_{d(c_j, x)}^\alpha(x, 0) \cap \Pi_1 = \emptyset) \mathbf{1}(d(c_j, x) \leq |x|) \right] dx \quad (17) \\ &= \lambda_1 \lambda_2 \int_{\mathbb{R}^2} \int_{\Delta_x} d(y, x) e^{-\lambda_2 \pi |x|^2 - \lambda_1 |O_{d(y, x)}^\alpha(x, 0)|} dy dx \quad (18) \end{aligned}$$

where  $\Delta_x = \{y \in \mathbb{R}^2 \mid \alpha|x| \leq d(y, x) \leq |x|\} = O_{|x|}^\alpha(x, 0)$ . Eq.(15) follows by the independence of  $\Pi_0$  and  $\Pi_1, \Pi_2$ . Eq.(16) follows by the refined Campbell's Theorem [20]. In Eq.(17) we have explicitly computed the void probability for the event associated with  $\Pi_2$  and used the independence of  $\Pi_1$  and  $\Pi_2$ . Finally in Eq.(18) we have applied Campbell's Theorem once more and explicitly computed the void probability of the remaining event. Note that for a sensor at position 0 and sink at location  $x$ , the energy cost function  $d(y, x)$  associated with using an intervening compressor  $y$  must satisfy  $\alpha|x| \leq d(y, x) \leq |x|$  hence the range of integration in  $\Delta_x$ . This completes the proof of (4). The right hand side of Eq.(18) can be rewritten as

$$\lambda_1 \lambda_2 \int_{\mathbb{R}^2} e^{-\lambda_2 \pi |x|^2} \left\{ \int_{\Delta_x} d(y, x) e^{-\lambda_1 |O_{d(y, x)}^\alpha(x, 0)|} dy \right\} dx. \quad (19)$$

Now let  $u = d(y, x)$  and note that the inner integration in (19) depends only on  $u$ , since  $x$  constant and  $|O_{d(y, x)}^\alpha(x, 0)| = |O_u^\alpha(x, 0)|$ . This permits a change of variables, and rewriting the inner integral in (19) as follows –see [17] for more details.

$$\begin{aligned} & \lambda_1 \lambda_2 \int_{\mathbb{R}^2} e^{-\lambda_2 \pi |x|^2} \int_{\alpha|x|}^{|x|} u e^{-\lambda_1 |O_u^\alpha(x, 0)|} \frac{\partial |O_u^\alpha(x, 0)|}{\partial u} du dx \\ &= \lambda_2 \int_{\mathbb{R}^2} e^{-\lambda_2 \pi |x|^2} \left\{ -|x| e^{-\lambda_1 f(\alpha)|x|^2} + \alpha|x| \right. \\ & \quad \left. + \int_{\alpha|x|}^{|x|} e^{-\lambda_1 |O_u^\alpha(x, 0)|} du \right\} dx. \quad (20) \end{aligned}$$

Now we use the following fact to bound the integration within the bracket of the right hand side of Eq. (20).

*Fact 1: The area of a Cartesian oval  $O_u^\alpha(x, 0)$ , when  $\alpha|x| \leq u \leq |x|$  and  $0 \leq \alpha \leq 1$ , is lower bounded by*

$$|O_u^\alpha(x, 0)| \geq \frac{f(\alpha)}{(1-\alpha)^2} (u - \alpha|x|)^2$$

with  $f(\alpha)$  as given by (2). The bound is tight for  $\alpha = 0$ . Thus (20) can be bounded as follows:

$$(20) \leq \lambda_2 \int_{\mathbb{R}^2} e^{-\lambda_2 \pi |x|^2} \left\{ -|x| e^{-\lambda_1 f(\alpha) |x|^2} + \alpha |x| + \int_{\alpha |x|}^{|x|} e^{-\lambda_1 \frac{f(\alpha)}{(1-\alpha)^2} (u-\alpha |x|)^2} du \right\} dx$$

$$= \frac{\alpha}{2\sqrt{\lambda_2}} + \frac{1-\alpha}{2\sqrt{\lambda_2 + \lambda_1 f(\alpha)/\pi}} - \frac{\lambda_2 \pi \Gamma(3/2)}{(\lambda_2 \pi + \lambda_1 f(\alpha))^{3/2}}.$$

This gives the desired upper bound  $\hat{G}_U$  as follows.

$$G_U = \frac{\lambda_0}{\lambda_2} \cdot \left\{ \frac{\lambda_2 \pi \Gamma(3/2)}{(\lambda_2 \pi + \lambda_1 f(\alpha))^{3/2}} \right.$$

avg. cost when the typical sensor is in a sink cell

$$+ \lambda_1 \lambda_2 \int_{\mathbb{R}^2} \int_{\Delta_x} d(y, x) e^{-\lambda_2 \pi |x|^2 - \lambda_1 |O_{d(y,x)}^\alpha(x,0)|} dy dx \left. \right\}$$

avg. cost when the typical sensor is in a compressor cell

$$\leq \frac{\lambda_0}{\lambda_2} \left\{ \frac{\alpha}{2\sqrt{\lambda_2}} + \frac{1-\alpha}{2\sqrt{\lambda_2 + \lambda_1 f(\alpha)/\pi}} \right\} = \hat{G}_U.$$

REFERENCES

[1] J.M. Kahn, R.H. Katz, and K.S. Pister, "Next century challenges: Mobile networking for smart dust," in *Proc. MOBICOM*, 1999.

[2] D. Estrin, R. Govindan, J. Heidmann, and S. Kumar, "Next century challenges: Scalable coordination in sensor networks," in *Proc. MOBICOM*, 1999.

[3] I.F. Akyldiz, W. Su, Y. Sankarasubermanian, and E. Cayirici, "A survey on sensor networks," *IEEE Communications Magazine*, vol. 40, no. 8, pp. 102–114, August 2002.

[4] A.J. Goldsmith and S.B. Wicker, "Design challenges for energy-constrained ad hoc wireless networks," *IEEE Wireless Communications*, vol. 9, no. 4, 2002.

[5] S. S. Pradhan, J. Kusuma, and K. Ramachandran, "Distributed compression in a dense microsensor network," *IEEE Signal Processing Magazine*, pp. 51–60, March 2002.

[6] A. Scaglione and S. D. Servetto, "On the interdependence of routing and data compression in multi-hop sensor networks," in *Proc. ACM Mobicom*, 2002.

[7] J. Chou, D. Petrovic, and K. Ramchandran, "A distributed and adaptive signal processing approach to reducing energy consumption in sensor networks," in *Proc. IEEE Infocom*, 2003.

[8] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in *Hawaii International Conference on System Sciences*, 2000.

[9] D. Petrovic, R.C. Shah, K. Ramchandran, and J. Rabaey, "Data funneling: Routing with aggregation and compression for wireless sensor networks," in *Proc. IEEE Sensor Network Protocols and Applications*, May 2003.

[10] S. Singh, M. Woo, and C. Raghavendra, "Power-aware routing in mobile ad hoc networks," in *Proc. ACM Mobicom*, October 2002.

[11] V. Rodoplu and T. H. Meng, "Minimum energy mobile wireless networks," *IEEE JSAC*, vol. 17, no. 8, pp. 1333–1344, 1999.

[12] B. Sadeghi, V. Kanodia, A. Sabharwal, and E. Knightly, "Opportunistic media access for multirate ad hoc networks," in *Proc. ACM Mobicom*, September 2002.

[13] E. Duarte-Melo, M. Liu, and A. Misra, "A computational approach to the joint design of distributed data compression and data dissemination in a data-gathering wireless sensor network," in *34th Allerton Conference on Communications and Control*, 2003.

[14] S. Bandyopadhyay and E. Coyle, "An energy efficient hierarchical clustering algorithm for wireless sensor networks," in *Proceedings IEEE Infocom*, 2003.

[15] T. M. Cover, "A proof of the data compression theorem of Slepian and Wolf for ergodic sources," *IEEE Trans. Information Theory*, vol. 22, pp. 226–228, 1975.

[16] J. Edmonds, "Submodular functions, matroids and certain polyhedra," in *Proc. Calgary Int. Conf. Combinatorial Structures and Applications*, 1970, pp. 69–87.

[17] S. J. Baek, G. de Veciana, and X. Su, "Minimizing energy consumption in large-scale sensor networks through distributed data compression and hierarchical aggregation," *Technical Report, Wireless Networking and Communication Group, University of Texas, Austin, TX*, 2004.

[18] F. Baccelli and S. Zuyev, "Poisson-voronoi spanning trees with applications to the optimization of communication networks," *INRIA Research Report No.3040*, Nov. 1996.

[19] J. Moller, "Random Johnson-Mehl tessellations," *Adv. Appl. Prob.*, vol. 24, pp. 814–844, 1992.

[20] D. Stoyan, W. S. Kendall, and J. Mecke, *Stochastic Geometry and its Applications*, J. Wiley & Sons, Chichester, 1995.

[21] S. D. Servetto and G. Barnechea, "Constrained random walks on random graphs: Routing algorithms for large scale wireless sensor networks," in *Proceedings of the 1st ACM International Workshop on Wireless Sensor Networks and Applications*, 2002.

[22] J.H. Chang and L. Tassiulas, "Maximum lifetime routing in wireless sensor networks," in *Proceedings of Advanced Telecommunications and Information Distribution Research Program, College Park, MD*, 2000.

[23] F. Baccelli, M. Klein, M. Lebourges, and S. Zuyev, "Stochastic geometry and architecture of communication networks," *J. Telecommunication Systems*, vol. 7, pp. 209–227, 1997.

[24] J. Moller, "Random tessellations in  $\mathbb{R}^d$ ," *Adv. Appl. Prob.*, vol. 21, pp. 37–73, 1989.



**Seung Jun Baek** received his B.S. degree from Seoul National University in 1998, and M.S. degree from University of Texas at Austin in 2002, all in Electrical and Computer Engineering. He is currently working toward the Ph.D. degree in University of Texas at Austin. His research interest includes wireless ad-hoc networks, sensor networks, Multiple-Input-Multiple-Output systems and network protocols.



**Gustavo de Veciana** (S'88-M'94-SM 2001) received his B.S., M.S, and Ph.D. in electrical engineering from the University of California at Berkeley in 1987, 1990, and 1993 respectively. He is currently a Professor at the Department of Electrical and Computer Engineering at the University of Texas at Austin. His research focuses on the design, analysis and control of telecommunication networks. Current interests include: measurement, modeling and performance evaluation; wireless and sensor networks; algorithms for computer aided design of reliable systems. Dr. de Veciana has been an editor for the IEEE/ACM Transactions on Networking. He is the recipient of a General Motors Foundation Centennial Fellowship in Electrical Engineering and a 1996 National Science Foundation CAREER Award, and co-recipient of the IEEE Bill McCalla Best ICCAD Paper Award for 2000.



**Xun Su** received his BSEE from University of Electronic Science and Technology of China in 1992, and his MSEE from Southeast University, China in 1995. He received his Ph.D. in Electrical Engineering from UT Austin in December 2002, with a focus on dynamic network routing algorithms. He is now a network engineer in California Institute of Technology. His research interests include network protocol design, wireless networking, peer to peer systems and network measurement. He is a member of IEEE and SIAM.