Power-Efficient Sensor Placement and Transmission Structure for Data Gathering under Distortion Constraints[1]

Deepak Ganesan [†] Răzvan Cristescu [†‡] Baltasar Beferull-Lozano [†‡†]

† Department of Computer Science, University of Massachusetts,
Amherst, MA 01003

†‡ Center for Mathematics of Information
Caltech 1200 E. California Blvd, Caltech 136-93 Pasadena, CA 91125

†‡† Audio-Visual Communications Laboratory (LCAV),
Swiss Federal Institute of Technology (EPFL), Lausanne CH-1015, Switzerland

dganesan@cs.umass.edu, razvanc@caltech.edu, Baltasar.Beferull@epfl.ch

We consider the joint optimization of sensor placement and transmission structure for data gathering, where a given number of nodes need to be placed in a field such that the sensed data can be reconstructed at a sink within specified distortion bounds while minimizing the energy consumed for communication. We assume that the nodes use either joint entropy coding based on explicit communication between sensor nodes, where coding is done when side information is available, or Slepian-Wolf coding where nodes have knowledge of network correlation statistics. We consider both maximum and average distortion bounds. We prove that this optimization is NP-complete since it involves an interplay between the spaces of possible transmission structures given radio reachability limitations, and feasible placements satisfying distortion bounds.

We address this problem by first looking at the simplified problem of optimal placement in the one-dimensional case. An analytical solution is derived for the case when there is a simple aggregation scheme, and numerical results are provided for the cases when joint entropy encoding is used. We use the insight from our 1-D analysis to extend our results to the 2-D case and compare it to typical uniform random placement and shortest-path tree. Our algorithm for two-dimensional placement and transmission structure provides two to three fold reduction in total power consumption and between one to two orders of magnitude reduction in bottleneck power consumption. We perform an exhaustive performance analysis of our scheme under varying correlation models and model parameters and demonstrate that the performance improvement is typical over a range of data correlation models and parameters. We also study the impact of performing computationally-efficient data conditioning over a local scope rather than the entire network. Finally, we extend our explicit placement results to a randomized placement scheme and show that such a scheme can be effective when deployment does not permit exact node placement.

Categories and Subject Descriptors: H.1.1 [**MODELS AND PRINCIPLES**]: Systems and Information Theory – Information theory; C.2.4 [**COMPUTER-COMMUNICATION NETWORKS**]: Distributed Systems – Distributed applications

General Terms: Algorithms, Performance, Design, Theory

Additional Key Words and Phrases: Sensor networks, Data gathering, Information theory, Sensor node placement, Energy efficiency, Sensing distortion

[1]This work was supported (in part) by the National Competence Center in Research on Mobile Information and Communications Systems (NCCR-MICS), supported by the Swiss NSF under grant number 5005-67322. Part of this work has been presented at the Symposium on Information Processing in Sensor Networks (IPSN '04).

1. INTRODUCTION

Wireless sensor networks are often envisaged to comprise thousands of nodes accomplishing a sensing task. Yet, the realities of economies of scale in manufacturing and the high cost of many sensors themselves mean that these nodes are currently significantly more expensive than predicted. Therefore, typical deployed networks (e.g. habitat monitoring [Hamilton]) comprise a few hundred of nodes, each with cost of a few hundreds of dollars. While we await a future with ubiquitous cheap sensor nodes, a problem that is both immediate and necessary is to accomplish the required tasks with a limited number of resource-constrained sensor nodes.

We consider the problem of deploying a finite number of sensor nodes in a geographic area, and choosing a communication structure among the nodes of the corresponding network. A single sink is responsible for gathering the sensor data, for storage or control purposes. Since sensor nodes have limited battery power, an important goal is to minimize the total power consumption of data gathering, while keeping the sensing distortion within specified bounds. An important characteristic of typical sensor networks, that can be exploited for reducing the power consumption, is that the data measured at nodes is correlated.

Several algorithms have been proposed for energy efficient data gathering [Lindsey et al. 2001; Heinzelman et al. 2000; Intanagonwiwat et al. 2000]. However, these works do not take into consideration the correlation in the data. Recent studies on the joint rate allocation and transmission structure optimization for sensor networks with correlated data can be found in [Goel and Estrin 2003; Cristescu et al. 2004]. In particular, the result of [Cristescu et al. 2004] is similar in that it considers optimal tree structures for data gathering that exploit correlation in the measured data. Our work adds new constraints to this problem by allowing node placement to be varied, thus introducing tough distortion constraints. This results in a more complex problem that both requires a different approach to solve and produces novel results. Node placement for optimal coverage is a well-studied (and difficult) problem (e.g. [Eidenbenz 2002]). The work of [Dasgupta et al. 2003] considers the problem of energy-efficient topology aware placement. That work does not exploit the correlation present in the data measured; also, the placement constraints considered in that work are rather event driven than related to the distortion of measurement. Recently, [Cheng et al. 2004] studied the problem of energy-optimized node placement in two-dimensions with the goal of network lifetime. Our work makes a significantly more comprehensive study of the placement problem, both under different coding schemes as well as under different correlation models. This work is an extended version of prior results that were published in [Ganesan et al. 2004].

A commonly used method for deploying sensor networks is the uniform random placement, since such a deployment is often the easiest and cheapest. However, we believe that there are compelling reasons for understanding the interactions between the node placement and the data and transmission structures, and the effect of these interactions on the efficiency of utilization. First, studying the impact of placements lets us understand if other easy-to-deploy configurations of sensor nodes can give important gains in power consumption. As we show in this paper, this is likely to be the case. Second, controlled

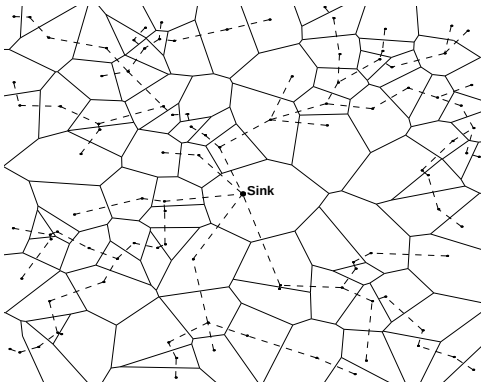

Fig. 1. The Voronoi cells (solid lines) represent the distortion in each cell. The tree structure (dotted lines) represents a possible transmission structure.

placement will be necessary for applications which have to deploy limited numbers of expensive nodes (such as seismic nodes which need high precision) and hence the location of sensors has to be optimized.

In this work, we are particularly interested in the relation between the data reconstruction distortion that results from the node placement, and the power requirements of data gathering from the sensors. An instance of the problem is shown in Fig. 1 where the Voronoi cells represent the distortion, and the dotted lines describe a possible tree structure to be used for data gathering.

More specifically, N nodes need to be deployed over a finite geographic region $\mathcal{A}$, whose two-dimensional area is A. Each of these nodes takes samples from a three dimensional random field $X(u, v, t)$, where (u, v) is the spatial location and t is the time-axis. Each sensor transmits periodically its sensed data, through multi-hop routing, to a sink located at the center of $\mathcal{A}$. The sink can reconstruct the sensor field in the region within specified maximum and average distortion bounds (D_{max} and D_{avg} respectively). We assume that sampling in the time domain is sufficiently high (above Nyquist frequency) for the sink to fully reconstruct the time-axis. Thus, we only need to consider the distortion in reconstructing snapshots along the spatial axes, $X(u, v)$.

To enable data gathering, the sensor nodes build a routing tree rooted at the sink, and transmit data along this tree. Note that there are situations where tree structures are not optimal, but for the sake of simplicity we will limit our study to data gathering trees. The data gathering procedure is periodic and originates at the leaves, proceeding iteratively towards the sink through multihop forwarding. We consider two kinds of coding schemes to exploit correlations in sensor data. The first is joint entropy coding with explicit communication, where at each iteration, a junction node receives data from its children, decodes the received data, jointly codes the decoded data with its own data, and forwards the encoded data to its parent on the tree. The second coding scheme is Slepian-Wolf [Slepian and Wolf 1973], where the correlated data generated at nodes can be coded with a total rate not exceeding the joint entropy, even without nodes explicitly communicating with each other (under some constraints on the rates, given by the so-called Slepian-Wolf region). Explicit communication and Slepian-Wolf coding offer different computation-communication

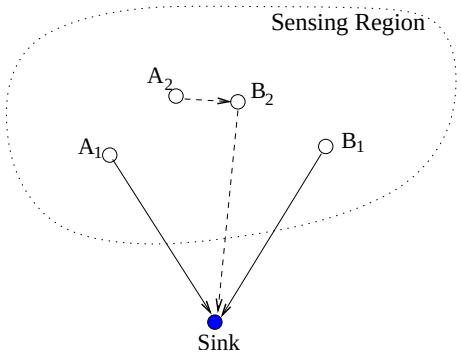

Fig. 2. Two possible placement and structure configurations of nodes A and B that satisfy specified sampling distortion bounds are shown. A_1 and B_1 transmit their data over shorter cumulative distances than A_2 and B_2 respectively, but B_2 can code its data with that of A_2 to exploit their high data correlation as a result of their proximity. Thus, determining the more power-efficient configuration among many possible ones that satisfy distortion bounds is difficult due to the interplay of placement and transmission structure.

tradeoffs. While coding is simpler in the case of joint entropy coding model with explicit communication, Slepian-Wolf typically offers larger communication gains [Cristescu et al. 2005].

We use a simple but relevant energy related cost function to study the interplay between placement (which determines locations of nodes) and transmission structure (which determines how sources on the tree are connected). Namely, the total cost of data gathering over the tree structure described above can be written as

$$\sum_{i=1}^{N} Rate(i) \times CommunicationCost(i) \tag{1}$$

where $Rate(i)$ is the total amount of data transmitted by node i, and *Communication-Cost(i)* is the per-bit transmission cost from node i to its parent on the tree. Since the data at nodes is correlated, the rate at node i, $Rate(i)$ depends on the particular set of sources that are in the sub-tree rooted at i, and on their locations. Similarly, the transmission cost per-bit, $CommunicationCost(i)$, depends on the identity of the particular node used as next hop to sink, and on its position. Thus, the total power consumption depends on both the placement and the transmission structure.

An illustration of this interplay is seen in Fig. 2, where even in a simple example with two nodes, it is not easy to determine the most power-efficient configuration. In terms of transmission distance, configuration 1 is better since A_1 and B_1 transmit each over smaller distance than A_2 and B_2, respectively. However, in configuration 2, A_2 and B_2 are closer to each other than nodes A_1 and B_1, hence they are likely to be stronger correlated. This correlation can be exploited by properly choosing the transmission structure ($A_2 \rightarrow B_2 \rightarrow sink$). This can significantly reduce the amount of data that node B_2 needs to generate, by using side information from node A in coding its data, thus reducing the total power consumption. On the other hand, nodes A_2 and B_2 cannot be placed arbitrarily

close to each other as this might violate the distortion constraints. In general, if the data is correlated, the shortest path tree (SPT) is not necessarily optimal as transmission structure [Cristescu et al. 2004].

The optimal solution involves searching through the spaces of all possible configurations that satisfy distortion bounds, and all possible transmission structures that are feasible given radio reachability limitations. While this combined optimization has not been considered in prior work, it is known that even a subset of our problem is NP-hard [Goel and Estrin 2003; Cristescu et al. 2004] (namely, the transmission structure optimization for a given placement). We extend these results to prove that the problem we consider in this work is NP-hard as well.

The rest of this paper is structured as follows. In Section 2, we formulate precisely our problem, and describe the sensing, communication, aggregation and data reconstruction models. In Section 3, we consider the one-dimensional variant of the problem, provide analytical solutions for a simplified aggregation model and analyze the joint encoding case numerically in detail. In Section 4, we perform a detailed study of the impact of data correlation model, model parameters and localized data conditioning. In Section 5, we extend the solutions from the one-dimensional case to the two-dimensional case and in Section 6, we show that it out-performs typical random placement approaches both in total and bottleneck power consumption. Finally, in Section 7, we conclude with a description of various applications of our work, and some interesting extensions.

To the best of our knowledge, this is the first study of the interaction between node placement under distortion constraints, transmission structure optimization, and rate allocation in the context of sensor networks that measure correlated data.

2. PROBLEM FORMULATION

We assume that we are given N nodes, that need to be placed in a two dimensional region $\mathcal{A}$, of area A. For simplicity, we assume that the region is circular, and has a radius L. The placement of nodes, $P = \{(x_i, y_i) \in \mathcal{A}, 1 \leq i \leq N\}$ is constrained by two distortion metrics, the maximum distortion D_{max}, defined as the maximum acceptable distortion at any point in $\mathcal{A}$, and the average distortion D_{avg}, defined as the distortion per unit area over $\mathcal{A}$.

2.1 Sensing Model

A frequently used sensing model is the Gaussian random field [Marco et al. 2003; Cristescu et al. 2004]. This model is suitable for analysis and can provide the essential intuition to solve the problem in practice. We assume that the field is a continuous-space two dimensional stationary random field $X(u, v)$, where u and v represent the geographic coordinates of points in the region $\mathcal{A}$. Without loss of generality, we assume that the random field has zero mean, that is, $\mu_X = E[X(u, v)] = 0 \quad \forall\, u, v$. We make simple assumptions about the nature of the random field, and make no assumptions on whether this field is band-limited or not. Let $\mathbf{R}_X(\tau_x, \tau_y)$ denote the covariance function associated to the random field $X(u, v)$. The correlation between two points (u_i, v_i) (position of node i), and (u_j, v_j) (position of node j) is given by $r_X(i, j) = \mathbf{R}_X(u_j - u_i, v_j - v_i)$.

We use two spatial data correlation models that are typically observed in spatial sensor datasets and used widely in spatial statistics [Cressie 1991]. The first model is one where correlation decreases exponentially with the distance between nodes, and the second model is one where the correlation decreases exponentially with the square of the

distance between nodes.

Markov Model. The correlation functions are $\mathbf{R}_X(\tau_x, \tau_y) = e^{-a\sqrt{\tau_x^2+\tau_y^2}}$, or in terms of Euclidean distance, $\mathbf{R}_X(d) = e^{-ad}$. Thus, $r_X(i,j) = e^{-ad_{ij}}$, where d_{ij} is the Euclidean distance between nodes i and j.

Square Decay Model. The correlation function for the square decay model is $\mathbf{R}_X(\tau_x, \tau_y) = e^{-a(\tau_x^2+\tau_y^2)}$, or in terms of Euclidean distance, $\mathbf{R}_X(d) = e^{-ad^2}$. Thus, $r_X(i,j) = e^{-ad_{ij}^2}$, where d_{ij} is the Euclidean distance between nodes i and j.

Having described the sensing model, we proceed to formulate the power consumption of a sensor node. The power consumption incurred depends on two factors: the distance from a transmitting node to its parent, and the aggregate amount of data transmitted over that distance.

2.2 Communication Model

For the transmission power, we use a standard transmission model that assumes that the power per bit for transmission over a wireless link is a function of the distance between the transmitter and receiver. We assume that there is an underlying transmission scheduling protocol (such as SMAC [Ye et al. 2002]) that schedules transmissions over a tree to avoids collisions. For the scope of this paper, we will ignore the protocol overhead resulting from creating the schedules, assuming that this overhead is small in comparison with the data size.

If the distance between a transmitter, i, and receiver, j, is d_{ij}, then the power is $E \propto d^{\kappa}$ where κ is called the path-loss exponent (typically $2 \leq \kappa \leq 4$ [Rappaport 1996][2]. In addition, each node has a maximum power at which it can transmit, which places a limit on the maximum transmission range, $d_{ij}^{\kappa} \leq E_{max}$. Thus

$$d_{ij} \leq C_{max} \tag{2}$$

where $C_{max} = (E_{max})^{\frac{1}{\kappa}}$ is the maximum communication radius of a sensor node.

Besides the power required for communication between adjacent sensor nodes, a second communication metric of importance is the quality of a wireless link, which is typically characterized by the packet throughput or alternately, packet-loss over a link. Packet-loss is typically dependent on the distance between nodes, with losses being higher for nodes with larger separation. This behavior is similar to the transmission power model described above, hence it can be easily integrated into our optimization. Therefore, we do not consider it explicitly in the rest of this paper.

The radio communication constraint overlaps with the maximum distortion constraint that we will discuss shortly. Both these constraints limit the maximum separation between nodes in the network.

Our model is simplistic to keep the optimization manageable for this paper. In practice, two additional factors need to be considered: (a) radios often have non-isotropic propagation, and, (b) radios adjust power levels in discrete steps rather than at arbitrarily fine granularity.

[2]In this work, we do not consider reception overhead, which would increase the cost of communication per-hop, and can only improve our results.

The above mentioned communication model determines the communication cost-per-bit. We now discuss how the total number of bits transmitted at each node is determined.

2.3 Aggregation Model

For our aggregation model, we assume that each node performs either joint entropy coding of the data coming from its corresponding sub-tree or Slepian-Wolf coding. We assume that each node quantizes its samples with an independent quantizer and all quantizers use the same quantization interval. In order to express the amount of data in bits resulting from first quantizing independently at each node and then performing joint entropy at a given node of all that data in its corresponding sub-tree, we will use the differential entropy as described in [Cover and Thomas 1991]. The differential entropy of a k-dimensional multivariate normal distribution $\mathcal{N}_k(\mu_X, \mathbf{R}_X)$ is:

$$h(\mathcal{N}_k(\mu_X, \mathbf{R}_X)) = \frac{1}{2}\log(2\pi e)^k \det(\mathbf{R}) \tag{3}$$

We approximate the joint discrete entropy associated to the quantized samples by assuming a high-resolution uniform scalar quantization with step-size Δ for all the nodes:

$$H(\mathcal{N}_k(\mu_X, \mathbf{R}_X)) \approx h(\mathcal{N}_k(\mu_X, \mathbf{R}_X)) - k\log\Delta \tag{4}$$

Thus, as $\Delta \to 0$, the distortion at each node $\to \Delta^2/12$ ([Cover and Thomas 1991]). Also, due to the fine quantization, we assume that the noise is uncorrelated with the signal.

2.3.1 *Joint Entropy Coding with Explicit Communication.* For any node i let T_i represent the set of nodes in the sub-tree rooted at node i. Node i receives encoded data from its children, first decodes it, and then jointly compresses it together with its own data (quantized samples). The total data rate sent from node i is approximately:

$$Rate(i) = H(\mathcal{N}_k(0, \mathbf{R}[T_i])) = \frac{1}{2}\log(2\pi e)^{|T_i|}\det(\mathbf{R}[T_i]) - |T_i|\log\Delta \tag{5}$$

where $\mathbf{R}[T_i]$ denotes the covariance matrix associated with the nodes in the sub-tree T_i, and $|T_i|$ represents the number of nodes of the sub-tree.

2.3.2 *Slepian-Wolf Coding.* To formulate the aggregation model for the Slepian-Wolf case, we use a result from [Cristescu et al. 2005]. This result shows that each node conditionally codes its data based on nodes "closer" than itself to the base. This is the counterpart of the joint entropy coding using explicit communication approach mentioned previously where each node conditions its data on nodes that are further from the base than itself. For any node i, let N_i denote the set of nodes closer to the base than node i. The incremental rate from a node i is then:

$$IncrementalRate(i) = H(i, N_i) - H(N_i) \tag{6}$$

The total data that a node i transmits is the cumulative sum of all data that is relayed through the node and the incremental data that node i generates. This total rate is:

$$Rate(i) \;=\; \sum_{j\in T_i} H(j, N_j) - H(N_j)$$

$$= \frac{1}{2}\log(2\pi e)^{|N_j|+1}\det(\mathbf{R}[j,N_j]) - (|N_j|+1)\log\Delta$$
$$- \frac{1}{2}\log(2\pi e)^{|N_j|}\det(\mathbf{R}[N_j]) + |N_j|\log\Delta \quad (7)$$

2.3.3 *Localized K-Hop Conditioning.* Equations (5) and (7) assumed that conditioning was performed over all nodes in the network. Thus, in the explicit communication case, an intermediate node performs joint entropy coding of all nodes that transmit data to it. Such coding, while efficient, is also very computation intensive. In localized conditioning, we exploit the fact that the correlation model is spatially localized and conditioning over nodes far away does not provide significant energy gains. Thus, a node conditions its data over data from other nearby nodes, but not over far away nodes.

The metric that we use for localized conditioning is the number of hops. Thus, a node conditions its data over all nodes that are within k-hops of itself, where k varies from 0 (no conditioning) to infinity (full conditioning).

Let $T_{i,k}$ represent the set of nodes in the subtree rooted at node i that are at most k hops from i. Then, the total data rate from i for the joint entropy coding using explicit communication case is:

$$H(\mathcal{N}_k(0,\mathbf{R}[T_{i,k}])) = \frac{1}{2}\log(2\pi e)^{|T_{i,k}|}\det(\mathbf{R}[T_{i,k}]) - |T_{i,k}|\log\Delta \quad (8)$$

Similarly, if $N_{i,k}$ represents the set of nodes that are closer than node i to the base station and within k hops of node i, then the total data rate when Slepian-Wolf coding is used is:

$$Rate(i) = \sum_{j\in T_i} H(j,N_{j,k}) - H(N_{j,k})$$
$$= \frac{1}{2}\log(2\pi e)^{|N_{j,k}|+1}\det(\mathbf{R}[j,N_{j,k}]) - (|N_{j,k}|+1)\log\Delta$$
$$- \frac{1}{2}\log(2\pi e)^{|N_{j,k}|}\det(\mathbf{R}[N_{j,k}]) - |N_{j,k}|\log\Delta$$

2.4 Data Reconstruction Model

We now describe the procedure used by the sink to reconstruct the entire continuous-space sensor field given the encoded data from a discrete set of sample points at the positions of the N nodes. The sink periodically receives quantized values $\hat{X}(u_1,v_1), \hat{X}(u_2,v_2)\ldots\hat{X}(u_N,v_N)$ from the N sensing nodes placed at points $(u_1,v_1),(u_2,v_2)\ldots$,$(u_N,v_N)$ respectively. In general, given these N quantized values, an interpolation procedure will result in a reconstruction, $\hat{X}(u,v)$, that gives the samples at any location (u,v) in the region $\mathcal{A}$. In this work, we use a nearest-neighbor reconstruction procedure, which, although very simple, helps us understand the complex interactions in our problem and focus on the power minimization issue. In future work, we plan to improve our results using better interpolation models[3].

Let V_i be the Voronoi cell corresponding to the sensor node i located at position (u_i,v_i) (which is the centroid of V_i). Then:

[3]Note that optimal reconstruction is difficult, because of the different issues of aliasing, non-uniform quantization, etc.

$$\begin{aligned}\hat{X}(u,v) &= \hat{X}(u_i,v_i) \text{ iff } ||(u,v)-(u_i,v_i)|| \\ &\leq ||(u,v)-(u_j,v_j)||, (\forall) j \neq i\end{aligned}$$

Given that the sink uses a nearest neighbor reconstruction procedure, we formulate the coverage and distortion constraints. The first constraint is a coverage constraint, which ensures that the set of Voronoi cells covers the region $\mathcal{A}$:

$$\bigcup_{i=1,..,N} V_i = A \tag{9}$$

To evaluate the maximum distortion, we use the fact that for an isotropic radially decreasing correlation model, the maximum distortion points are located along the boundaries of the Voronoi cells. By definition, all Voronoi cells are convex because of the property of minimum distance decoding, therefore, the furthest points in each Voronoi cell are the corners of the cell. The maximum distortion constraint controls the distance from these furthest points in each cell to the centroid of the corresponding Voronoi cell. Thus, for any point (u,v) in region $\mathcal{A}$, the distortion of reconstruction when it is assigned the same value as the nearest sampled point (u_i,v_i) is:

$$\begin{aligned}\text{Distortion}(u,v) &= MSE(u,v) \\ &= E[(\hat{X}(u_i,v_i)-X(u,v))^2] \leq D_{max}\end{aligned}$$

with i sensor being closest to (u,v). Note that the error is computed between the quantized version of the closest sample given by $\hat{X}(u,v)$ which is received at the sink, and the actual unquantized random variable, $X(u,v)$.

For instance, using the Markovian correlation model in Section 2.1, we obtain the distortion between the unquantized random variables at the two points:

$$\begin{aligned}MSE(u,v) &= E[(\hat{X}(u,v)-X(u,v))^2] \\ &= E[(\hat{X}(u_i,v_i)-X(u,v))^2] \\ &= E[(X(u_i,v_i)+n_Q(u_i,v_i))^2]+E[X(u,v)^2] \\ &\quad -2E[(X(u_i,v_i)+n_Q(u_i,v_i))X(u,v)] \\ &= E[X(u_i,v_i)^2]+E[X(u,v)^2]-2E[X(u_i,v_i)X(u,v)] \\ &= \sigma^2+\sigma^2-2\sigma^2 e^{-ad_{ij}} \\ &= 2\sigma^2(1-e^{-ad_{ij}}) \leq D_{max}\end{aligned} \tag{10}$$

where n_Q is the quantization noise between the quantized random variable and the original random variable, which is assumed to be small due to fine quantization. As expected, the above equation shows that the mean square error (MSE) is a concave and monotonically increasing function of the distance between the location (u,v) and the closest sample point. Therefore, the maximum distortion bounds the distance between any point in the region $\mathcal{A}$, and the nearest sample point. Thus, for the particular model that we consider, the maximum allowed distance R_{max}, is:

$$R_{max} = -\frac{1}{a}\log\left(1 - \frac{D_{max}}{2\sigma^2}\right) \tag{11}$$

The average distortion constraint, defined as the mean square error in data reconstruction over the entire region $MSE(\mathcal{A})$, can be computed by integrating $MSE(u,v)$ over $\mathcal{A}$.

$$\begin{aligned} \mathrm{A}vgDistortion(\mathcal{A}) &= MSE(\mathcal{A}) \\ &= \frac{1}{A}\int_{\mathcal{A}} MSE(u,v)dudv \leq D_{avg} \end{aligned} \tag{12}$$

where $MSE(u,v)$ is calculated as shown in (10). Then, for N sensors, each with corresponding Voronoi cell V_i,

$$\begin{aligned} MSE(\mathcal{A}) &= \frac{1}{A}\sum_{i=1}^{N} MSE(V_i) \\ &= \frac{1}{A}\sum_{i=1}^{N}\int_{V_i} MSE(u,v)dudv \\ &= \frac{1}{A}\sum_{i=1}^{N}\int_{V_i} E[(\hat{X}(u,v) - X(u,v))]dudv \leq D_{avg} \end{aligned} \tag{13}$$

2.5 Objective Function

We state now formally our objective to minimize the total power cost, given the constraints described so far. Namely, our problem is to find a placement P of nodes, $|P| = N$, and a tree ST rooted in the sink, that spans the nodes in P, such that to

$$\text{Minimize (1)} \quad \text{under constraints (2), (9), (10), (12).} \tag{14}$$

We study (14) for both coding strategies (5) and (7).

3. OPTIMAL PLACEMENT AND STRUCTURE IN THE ONE-DIMENSIONAL CASE

Although our final goal is to solve the problem in the two-dimensional case, the one-dimensional case is significantly more tractable since, as we will show shortly, the transmission structure optimization is trivial, and it is possible to understand the placement problem in isolation.

We adapt the problem statement in Section 2 for the one-dimensional instance. In this case, let $X(s)$, $0 \leq s \leq L$, represent the measured random field along a one-dimensional line of length L, as shown in Fig. 3. Since all transmission terminates at the sink, instances of the problem where the sink is between nodes (not at the corner of a line) can be split into two independent optimizations, one each for nodes on either side of the sink with the sink at the corner. Further, due to symmetry, it does not matter which end of the line the sink is placed. Therefore, the N nodes $1, 2, .., N$ are placed in sequence along the line with the sink at the left corner such that node 1 is closest to the sink and node N is the furthest. We denote the distance between node i and node $i-1$ by r_i; $\{r_i\}_{i=1}^{N}$ represents the set of unknowns in the one-dimensional optimization.

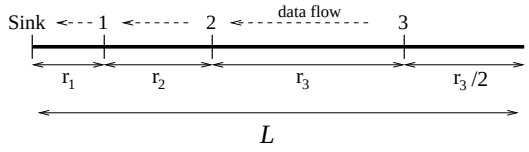

Fig. 3. One-dimensional node placement.

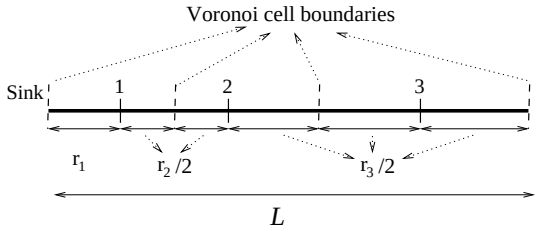

Fig. 4. Voronoi cells and maximum distortion distances for linear placement.

3.1 One-Dimensional Constraints

For the one-dimensional setting, the coverage, transmission and distortion constraints can be readily formulated similar to (9), (10) and (12). The Voronoi cells for the one-dimensional problem are the mid-points between adjacent pairs of sample points (nodes) as shown in Fig. 4. Without loss of generality, we use a boundary extension of the Voronoi cell for the last node, r_N. Thus, the coverage constraint can be rewritten as:

$$\text{Coverage Constraint: } \sum_{i=1}^{N} r_i + r_N/2 = L. \tag{15}$$

The communication constraint limits the separation between nodes as shown in (2). Thus,

$$r_i \leq C_{max} \tag{16}$$

As described in (10), the maximum distortion constrain restricts the maximum distance from any node to the edge of the nearest neighbor cell. Hence,

$$r_1 \leq R_{max} \qquad \frac{r_i}{2} \leq R_{max}\ ; \quad (\forall) \quad 2 \leq i \leq N \tag{17}$$

As can be seen, the maximum distortion constraint is very similar to the communication constraint.

Similar to (12), the average distortion (defined as the distortion per unit length) is given by:

$$\begin{aligned} MSE(L) &= \frac{1}{L}\int_L MSE(u)du \\ &= \frac{1}{L}\sum_1^N \int_{V_i} E\left[(\hat{X}(u) - X(u))\right] du \end{aligned}$$

$$= \frac{1}{L}\left[\int_0^{r_1} 2\sigma^2(1-e^{-ad}) + 2\sum_2^N \int_0^{\frac{r_i}{2}} 2\sigma^2(1-e^{-ad}) + \int_0^{\frac{r_n}{2}} 2\sigma^2(1-e^{-ad})\right]$$

$$= 2 - \frac{2}{La}\left(2N - e^{-ar_1} - 2\sum_2^N e^{-ar_i/2} - e^{-ar_N/2}\right) \quad (18)$$

We analyze the joint placement-structure optimization in two steps. First, we show that the optimal structure is simple shortest path routing. We then proceed to optimizing the placement for different choices of aggregation functions.

3.2 Optimal Structure is Shortest Path

Proposition 1: In a one-dimensional sensor network where there is a single sink and joint entropy coding is used at each hop, shortest path communication is optimal in terms of minimizing total energy.

This proposition can be easily proven since power increases super-linearly with distance and joint entropy increases with the number of nodes aggregated. If node i transmits its data to a node j where $j > i$ (i.e. j is further from the sink than i), the data from j must be eventually routed through i to minimize power consumption. This results from the fact that power per-bit increases super-linearly with distance (since κ is between 2 and 4), hence it is always better to multi-hop through as many intermediate hops as available [Pottie and Kaiser 2000]. Thus, to minimize power consumption, the aggregate data from j must be routed through i. On the other hand, joint entropy coding is a monotonically increasing function of the number of sources, hence, if node i transmits its data to a node j where $j > i$, the jointly coded data at j is larger than the amount of data when i did not transmit to j, and this consumes more power to transmit to the sink.

3.3 Optimizing Placement for 1-D Transmission

Given that the transmission structure is shortest path forwarding from nodes towards the sink, we now reformulate the placement problem for the two explicit communication and Slepian-Wolf coding strategies.

3.3.1 *Explicit Communication.*

$$\{r_i\}_{i=1}^N = \arg\min_{r_i} \sum_{i=1}^N H(X_i, X_{i+1}, .., X_N) r_i^\kappa$$

$$= \arg\min_{r_i} \sum_{i=1}^N \left[\frac{1}{2}\log(2\pi e)^{(N-i+1)} \det(\mathbf{R}[X_i..X_N]) - (N-i+1)\log\Delta\right] r_i^\kappa \quad (19)$$

under coverage (15) and distortion constraints (17),(18).

3.3.2 *Slepian-Wolf Coding.*

For the Slepian-Wolf coding in the 1D case, we first derive the total rate from a node i along the line based on Equation 7.

$$
\begin{aligned}
Rate(i) &= (\sum_{j=i}^{N}(H(X_1,...,X_j) - H(X_1,...,X_{j-1}))) \\
&= (H(X_1,...,X_N) - H(X_1,...,X_{i-1})) \\
&= \left[\frac{1}{2}\log(2\pi e)^{(N)}\det(\mathbf{R}[X_1,...,X_N])\right] - N\log\Delta \\
&- \left[\frac{1}{2}\log(2\pi e)^{(i-1)}\det(\mathbf{R}[X_1,...,X_{i-1}])\right] + (i-1)\log\Delta \qquad (20)
\end{aligned}
$$

The placement problem can then be formulated as:

$$
\begin{aligned}
&\arg\min_{r_i}\textstyle\sum_{i=1}^{N} Rate(i).r_i^{\kappa} \\
&= \arg\min_{r_i}(\left[\tfrac{1}{2}\log(2\pi e)^{(N)}\det(\mathbf{R}[X_1,...,X_N])\right] - N\log\Delta \\
&- \left[\tfrac{1}{2}\log(2\pi e)^{(i-1)}\det(\mathbf{R}[X_1,...,X_{i-1}])\right] + (i-1)\log\Delta)r_i^{\kappa} \qquad (21)
\end{aligned}
$$

It is hard to solve (19) and (21) analytically for an arbitrary correlation structure since the correlation structure depends implicitely on the inter-node distances. We thus first obtain a closed-form solution for a simplified scenario, where we assume zero correlation between data sampled at different nodes. In this case, $\mathbf{R}[T_i]$ is diagonal $(\forall)i \in \{1...N\}$.

3.3.3 *Analytical Solution for Independent Data at Nodes.* In this case, the optimization in (19) and (21) reduces to minimizing $\sum_{i=1}^{N}\left[(N-i+1)(\frac{1}{2}\log(2\pi e) - \log\Delta)\right] r_i^{\kappa}$ since $\det(\mathbf{R}[T_i])$ is unity when nodes are uncorrelated and the variance of the random process is one. Dropping the constant scaling factor that does not impact the minimization, we get:

$$
\{r_i\}_{i=1}^{N} = \arg\min_{r_i}\sum_{i=1}^{N}(N-i+1)r_i^{\kappa} \qquad (22)
$$

As an example, suppose in Fig. 3 the samples at nodes 1, 2 and 3 are uncorrelated, and each of the nodes has one unit of data to transmit to the sink. In this case, node 3 transmits one unit of data, node 2 transmits two units (its own unit + one forwarded unit), and node 1 transmits three units. Let us now see the impact of each constraint in the above optimization.

First, we consider the optimization in (22) when only the coverage constraint (15) is active. Using a Lagrangian multiplier, we obtain

$$
\{r_i\}_{i=1}^{N} = \arg\min_{r_i}\sum_{i=1}^{N}(N-i+1)r_i^{\kappa} - \lambda\left[(\sum_{i=1}^{N} r_i) + r_N/2\right] \qquad (23)
$$

By solving this Lagrangian optimization using partial derivatives, we obtain:

$$
r_i = \left[\frac{\lambda}{\kappa(N-i+1)}\right]^{\frac{1}{\kappa-1}}, \quad (\forall)\ 1 \le i \le N-1;
$$

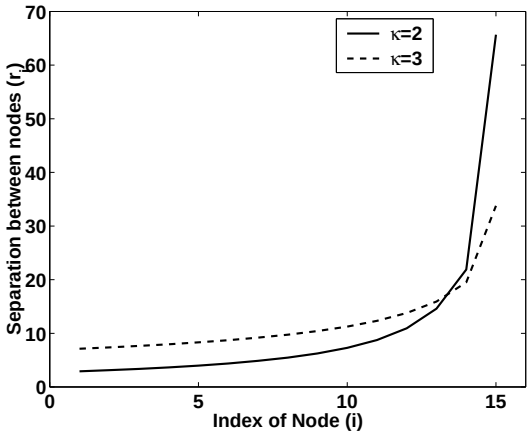

Fig. 5. Placement for different values of pathloss exponent (κ). As κ increases, placements become more uniform, since communication power dominates differences in aggregated data size.

$$r_N = \left[\frac{3\lambda}{2\kappa}\right]^{\frac{1}{\kappa-1}}$$

$$where\ \lambda = \left[\frac{L}{\sum_1^{N-1}(\frac{i=1}{\kappa(N-i+1)})^{\frac{1}{\kappa-1}} + \frac{3}{2}(\frac{3}{2\kappa})^{\frac{1}{\kappa-1}}}\right]^{\kappa-1}$$

Fig. 5 shows the optimal $\{r_i\}_{i=1}^N$ in a placement with $N = 15$ nodes over a line of length $L = 200$, for the case of quadratic and cubic path-loss exponents, $\kappa = 2$ and 3. As expected, nodes further from the sink transmit smaller amounts of data over longer distance than closer ones, which need to transmit larger amounts of aggregated data. However, the impact of increasing data load for nodes closer to the sink is balanced by the effect of the path-loss exponent κ, since communication power increases super-linearly with distance. Thus, the optimal choice reflects a balance between these two opposing factors. As the path-loss exponent is increased from $\kappa = 2$ to $\kappa = 3$, the communication overhead dominates, hence the spacing between nodes becomes more uniform.

The maximum and average distortion bounds impact on placement in significantly different ways as shown in Fig. 6. The maximum distortion constraint places a ceiling on the maximum separation between nodes, whereas the average distortion constraint reduces the mean error by making cells more equally sized.

4. PERFORMANCE OF 1D PLACEMENT

We evaluate the performance of our optimal one-dimensional placement for the two coding models by comparing its power consumption to that of a commonly considered uniform placement. Two performance metrics are considered:

—**Total Power Gain**: This metric measures the ratio of total power consumption for a regularly spaced placement to our optimal 1D placement.

—**Bottleneck Power Gain**: This metric measures the ratio in power consumption for the bottleneck node between the regular and optimized placement. Our problem formulation

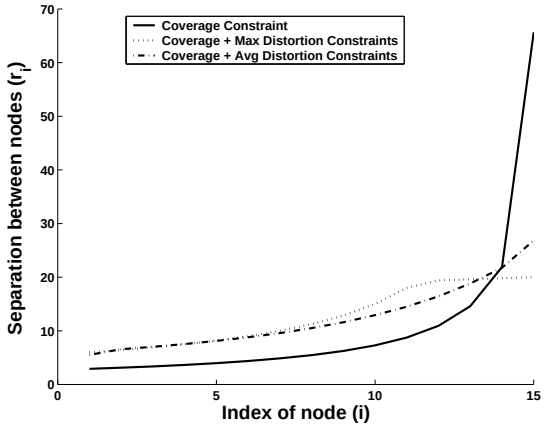

Fig. 6. Impact of distortion constraints. The R_{max} constraint places a ceiling on the maximum r_i whereas the D_{avg} constraint equalizes r_is to reduce the average distortion.

Pathloss Exp(κ)	R_{max}		N		L		D_{avg}	
	Small(10)	Large(20)	Small(10)	Large(20)	Small(100)	Large(300)	Small(0.05)	Large(0.15)
$\kappa = 2$	1.3	1.8	1.3	1.8	1.9	1.3	1.5	1.6
$\kappa = 3$	1.3	1.5	1.2	1.6	1.5	1.3	1.4	1.5

Table I. Gain of optimal placement over uniform placement for different settings.

optimizes for the *total* power consumption in data-gathering, but in a practical scenario, metrics such as network lifetime are likely to be as important. A commonly used metric for network lifetime is the time at which the first node dies, in other words, what is the power consumption of the bottleneck node. For instance, in a typical sensor network, the bottleneck node is the one that is closest to the sink, since it forwards a large amount of traffic.

We study the relative performance for two values (a relatively small one and a relatively large one, given our model) for each variable. The MATLAB optimization toolbox is used for numerically finding the optimal solution of (19) and (21) using the sensing model in Section 2.1 with parameter $a = 0.001$ (high correlation).

4.1 Performance of Joint Coding Case with Explicit Communication

The placement of nodes for the joint coding case with explicit communication has essentially the same behavior as that for the zero correlation case. This behavior results from the continuous and slowly decaying nature of our correlation model.

For a reasonable choice of distortion and network parameters, we see that even for low number of nodes we get a factor of 1.2 to 2 benefit over the uniform spacing case. For a larger network, these gains increase. Among other parameters, increasing R_{max} increases the flexibility in placement, hence it makes possible to further minimize power consumption. Increasing L for fixed R_{max} has the opposite effect, since the feasible placement region reduces. Increasing the number of nodes N (for fixed R_{max}), increases both the correlation between nodes (hence the aggregation benefit), and reduces the average per-

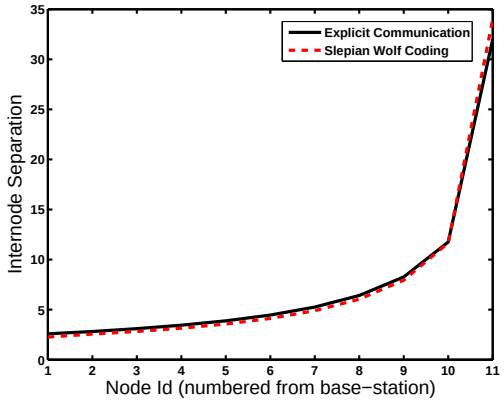

Fig. 7. The optimal placement for explicit communication and Slepian-Wolf coding are very similar. Slepian-Wolf coding results in slightly greater spacing for nodes farther from the base station and correspondingly closer for nearer ones. This is because farther nodes send less data in the Slepian-Wolf case than in the explicit communication case

hop distance for multihop transmission, thus, power consumption reduces.

4.2 Performance of Slepian-Wolf Coding

The performance in the Slepian-Wolf coding case is even better than that for explicit communication case as expected. Fig. 8 shows the percentage gain that the Slepian-Wolf 1D placement solution provides over joint entropy coding with explicit communication. As can be seen, Slepian-wolf coding is better by about 10-15% with gains reducing with more nodes being placed along a line. In more highly correlated settings, gains can be expected to be larger.

Fig. 7 shows a comparison of one instance of optimal placement for the Slepian-Wolf coding case with explicit communication. While the general trend is very similar, Slepian Wolf results in more irregular placements where the spacing between nodes farther from the base-station is larger than in the explicit communication case and the spacing between nodes closer to the base-station is less than in explicit communication. The intuition for this solution is that in the Slepian-Wolf Coding, the farther nodes transmit less data than their counterpart in the explicit coding case. The optimal solution thus spaces them farther apart.

4.3 Performance of K-Hop Conditioning

In the previous performance studies, we assumed that conditioning was performed over all nodes in the network. Thus, in the explicit communication case, an intermediate node performs joint entropy coding of all nodes that transmit data to it. Such coding, while efficient, is also very computation intensive. One technique to reduce the computational complexity is to condition over a limited scope rather than the entire network. The impact of localized conditioning is shown in Fig. 9. As can be seen, the power reduces with increased conditioning, but the marginal benefit reduces with each farther hop. Since data correlation decay over distance, the marginal power reduction due to condition over farther

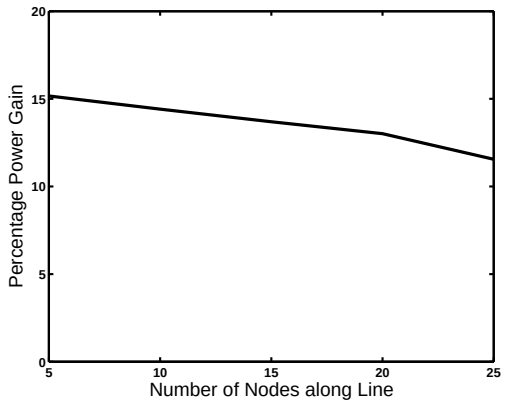

Fig. 8. Percentage gain offered by Slepian-Wolf coding over Joint Entropy coding with explicit communication.

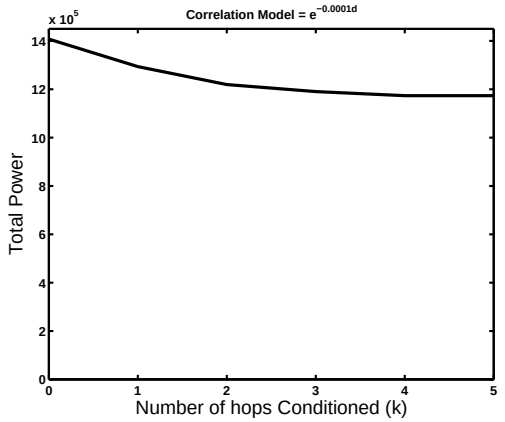

Fig. 9. As the number of hops over which the signal is conditioned increases, the marginal decrease in power consumption reduces.

nodes may not outweigh the significant additional computation required.

5. TWO-DIMENSIONAL PLACEMENT AND STRUCTURE

While the one-dimensional problem instance can be well-understood since node placement optimization separates from transmission structure optimization, the two-dimensional case is significantly more complex. In this section, we prove that the joint optimization of placement and structure in the two-dimensional case is NP-complete, and describe an approximation algorithms based on intuition from the solution in the 1D case.

5.1 Complexity

The proof of NP-completeness for the two-dimensional problem follows directly as a reduction from the problem of correlated data gathering with explicit communication, which

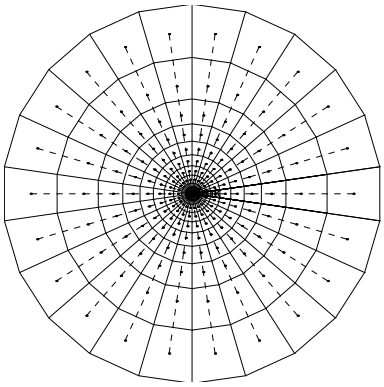

Fig. 10. The Voronoi cells (solid) and transmission structure (dashed) for a wheel placement, for data-gathering at a sink located at the center of the circular region $\mathcal{A}$.

is known to be NP-complete [Cristescu et al. 2004]. Namely, when the location of nodes is fixed, the problem of optimizing the transmission structure for power efficient data gathering is NP-complete. We show that the joint optimization problem of placement and transmission structure is NP-hard by proving that the decision version is NP-complete:

Proposition 2: For an arbitrary network and a given integer M, it is NP-complete to decide if there is any placement and transmission structure for which the cost in (1) is smaller than M, under constraints in (14).

Proof: If the aggregation function at nodes is concave and dependent on the number of nodes that relay via that node, then the optimization problem includes the Steiner tree problem; moreover, the problem is NP-complete also when the aggregation function is known [5, 16]. If the aggregation function at a node depends also on the transmission structure among the nodes that relay via that node, then, even for very simple settings, the problem remains NP-complete [6]. Thus, for every instance of the correlated data gathering problem with explicit communication, our reduction consists in assigning strong distortion constraints at each node, such that the position of the node that fulfills the distortion constraints becomes fixed. For such distortion constraints, the optimization problem is thus reduced to finding the optimal transmission structure, which is NP-complete [Cristescu et al. 2004].

5.2 Placement Strategy

Our two-dimensional placement strategy replicates the linear placement along a wheel structure as shown in Fig. 10. The wheel comprises n_{spoke} spokes where each spoke has n_{radial} nodes placed along it. Each node transmits data using shortest path forwarding along the spoke on which it is placed. Note that the shortest path might not be always optimal for explicit communication coding [Cristescu et al. 2004], however we restrict to such simple gathering trees, which can be constructed distributedly in polynomial time. We study in more detail the placement problem; the study of alternative transmission structures is subject of further work.

Fig. 10 shows both the transmission structure and the Voronoi cells for such a placement. Besides being analytically tractable, the wheel structure captures the essential behavior that

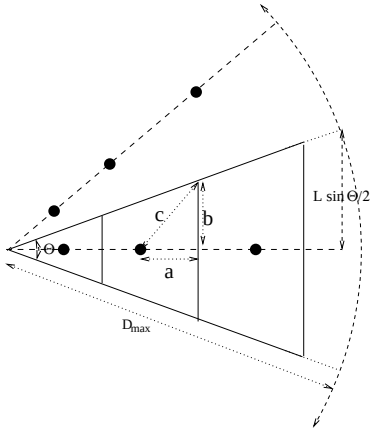

Fig. 11. Deriving the R_{max} constraint in two-dimensional placement. For any of the Voronoi cells, the furthest points are the corners of the cell.

we would want from an efficient two-dimensional placement and transmission structure. The network is dense closer to the sink where the data load is higher, and sparse further away from the sink.

While the two-dimensional placement is simple once we have decided that the placement and transmission structure is along a wheel, many questions remain to be solved. How do we place the n_{radial} nodes along each spoke such that the distortion bounds are not violated over the entire two-dimensional area, $\mathcal{A}$? Given N nodes, how many nodes, n_{radial}, do we place along each spoke and how many spokes, n_{spoke}, do we place angularly over the wheel? A lookahead into the results in Fig. 12 suggests that performance gains are not only sensitive to the choices of n_{spoke} and n_{radial}, but there is a non-obvious choice of these parameters that provides maximum power benefit.

Note that, even if we propose a deterministic placement strategy, our approach provides meaningful insight into the design of random radial distributions for nodes placement in an arbitrary area in practical scenarios.

5.2.1 *Maximum distortion bound.* We will translate the maximum distortion constraint for the two-dimensional case, to a one-dimensional bound that we can solve using the technique described in Section 3. Consider the Voronoi cells for a single spoke as shown in Fig. 11. As discussed in Section 2.4, the maximum distortion corresponds to the distortion at the point that is furthest from its nearest sampled point. By definition, such a point should lie on the Delaunay triangulation of the sample points. Since each Voronoi cell is convex, this point lies on one of the corners of the cell. Due to radial symmetry, the Voronoi cells for nodes on each spoke are identical, hence, it suffices to consider the maximum distortion bound on cells corresponding to any one of the spokes.

Consider the triangle with sides a, b and c that is formed between any sampling point and one of the corners of its Voronoi cell (see Fig. 11). The maximum distortion constraint is satisfied if $c \leq R_{max}$ (11). Since the radius of the region $\mathcal{A}$ is L, the angle between any two spokes is $\theta = \left(\frac{2\pi}{n_{spoke}}\right)$, and $b \leq L \sin\left(\frac{\theta}{2}\right)$, then it follows that a sufficient condition for $c \leq R_{max}$ is $a \leq \sqrt{R_{max}^2 - L^2 \sin^2\left(\frac{\theta}{2}\right)}$.

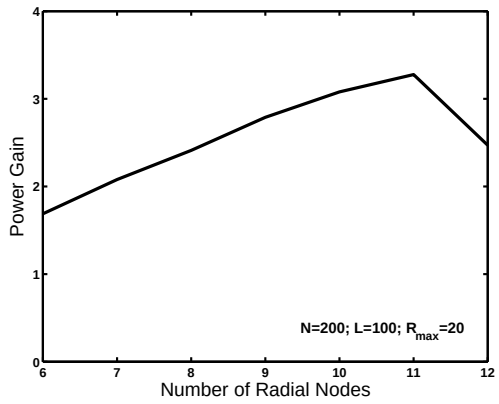

Fig. 12. As the number of nodes per-spoke (n_{radial}) is increased, gains initially improve but eventually $\bar{R}_{max}$ is too constrained, hence gains reduce.

Thus, in order to have the two-dimensional maximum distortion distance of R_{max} satisfied, it suffices to place nodes along each spoke such that the one-dimensional distortion distance along each line is bounded by:

$$\bar{R}_{max} = \sqrt{R_{max}^2 - L^2 \sin^2\left(\frac{\pi}{n_{spoke}}\right)} \tag{24}$$

5.3 Choosing the Number of Spokes and Nodes per Spoke

The tradeoff involved in finding the optimal choice of n_{spoke} and n_{radial} can be understood from (24). From the one-dimensional analysis, we know that separately increasing either n_{radial} or $\bar{R}_{max}$, and keeping the other constant, can reduce power consumption. However, in this case, the two parameters have opposite effects on each other. For instance, if n_{radial} is increased, n_{spoke} decreases and, therefore, so does $\bar{R}_{max}$ (24). This interaction is illustrated in Fig. 12.

How do we obtain a good choice of n_{radial} and n_{spoke}? While an exact solution for determining the optimal choice of wheel placement parameters is hard to find, an intuitive and effective approximation is to find the placement that maximizes $n_{spoke} \times \bar{R}_{max}$. Our approximation algorithm performs within 10% of the optimal (computed through exhaustive search) for configurations that we have tested.

5.4 Randomized Approximations of Explicit Placement

In this section we study the effects due to imperfect node placement on our optimized location strategies.

In Fig. 13, the placement of nodes is generated such that the distance from the nodes to the sink follows a probability distribution derived from our *optimal* placement, namely $cdf(d) = N(d)/N$, where d is the distance from the node to the sink, $N(d)$ is the number of nodes at a distance smaller than d from the sink in the optimal placement and $cdf(d)$ is the cumulative distribution function that we consider. We plot the ratio of costs between

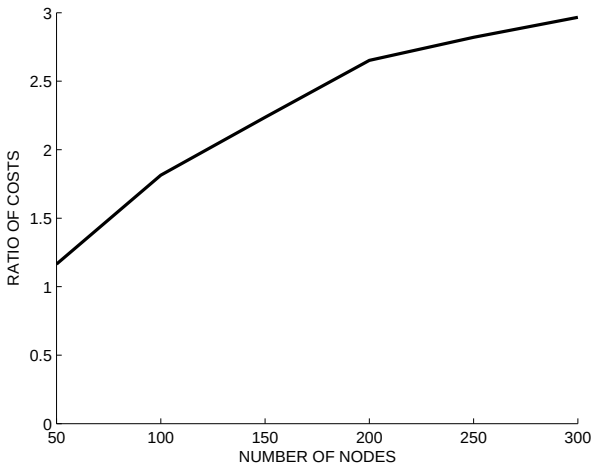

Fig. 13. The ratio of costs between randomized and optimal placement, for various network sizes N.

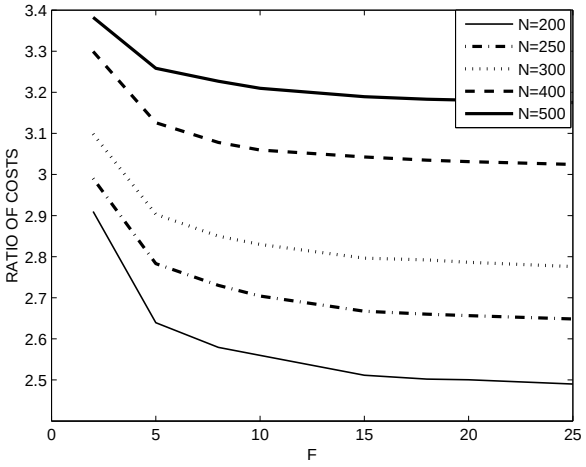

Fig. 14. The ratio of costs between optimized and uniform placement, for various values of the perturbation parameter F.

the probabilistic placement with $cdf(d)$ and the true optimal placement.

In Fig. 14, we model the jitter: the nodes are not in the optimal positions, but rather perturbed ones: $d_i = d_i + (d_{i+1} - d_{i-1})/F * w$, where d_i is the distance from node i to the sink, and w is normal distributed jitter noise $\mathcal{N}(0, 1)$ and F a parameter modelling the variance of the perturbation. We plot the ratio of costs between uniform placement and the optimal placement, both perturbed with the same F. Note that the distortion constraints are not guaranteed any longer, however given the values of F they are exceeded to a small extent.

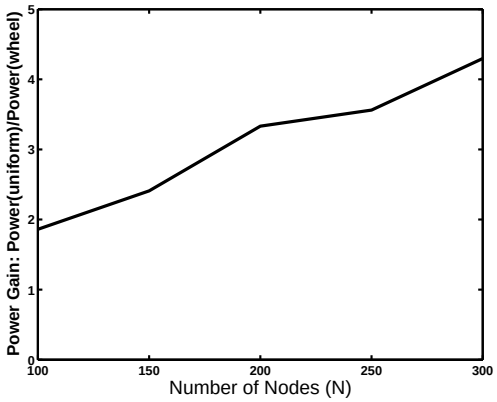

Fig. 15. As the number of nodes placed in the network is increased, the gains increase.

6. PERFORMANCE EVALUATION OF TWO-DIMENSIONAL PLACEMENT AND STRUCTURE

We use two metrics to evaluate the performance of our scheme, total power gain and bottleneck power gain.

6.1 Total Power Gain

Our first evaluation metric is the gain that the optimized two-dimensional placement provides over a commonly employed uniformly random node placement. Fig. 15 plots the improvement of an efficient wheel placement (as determined by the above metric) over the uniformly random placement in the circular area $\mathcal{A}$, where the transmission structure is a shortest path tree. When the number of nodes in the network, N is low, the performance gains are low for both coding schemes, since the distortion bounds provide less flexibility to optimize placement. As N increases, these gains increase up to a factor of 5.

The power gains observed using Slepian-Wolf coding follow the same trend as Explicit Communication but are better by a factor of 50%. The improvement depends on the choice of correlation parameters; in a highly correlated environment with many sensors, more gains can be expected.

Thus, not only does our wheel placement consistently out-perform a random placement with shortest path trees, it can potentially provide an order of magnitude improvement if sufficient flexibility is allowed in terms of number of nodes and distortion bounds.

6.2 Bottleneck Power Gain

A commonly used metric for network lifetime is the time at which the first node dies, in other words, what is the power consumption of the bottleneck node. For instance, in a typical sensor network, the bottleneck node is the one that is closest to the sink, since it forwards a large amount of traffic.

Fig. 16 shows that the power consumption of the bottleneck node in the optimal placement is two orders of magnitude lower than in the uniform random placement. The gains are roughly similar in magnitude irrespective of the coding scheme being used. The bottleneck power consumption when Slepian-Wolf coding is used is slightly lower than the

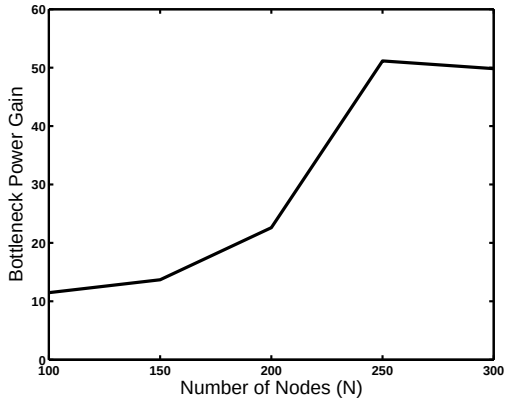

Fig. 16. The power gains at the bottleneck node are very large, with 50 times improvement for large N.

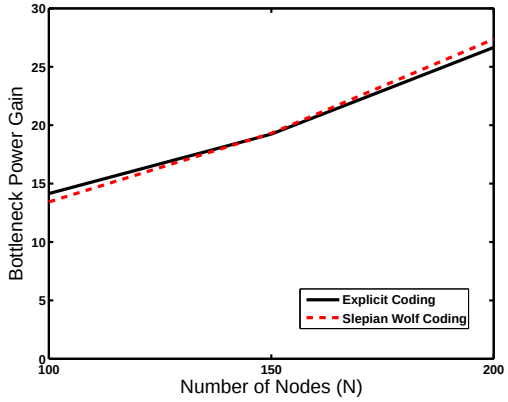

Fig. 17. The power gains for the Slepian-Wolf coding case are similar to those for explicit communication.

explicit communication case for low total number of nodes (N) and high for greater number of nodes.

The gains in bottleneck power consumption are due to the fact that our optimized scheme places nodes progressively closer to each other as they aggregate more data. This results in nodes that transmit more data being compensated by having to transmit over a shorter distance. Thus, huge gains in network lifetime can be expected from our optimized placement.

A unique feature of our optimized placement is that the bottleneck node is the *farthest* node from the base-station rather than the closest one as is the case with typical random or uniform placement schemes. This observation is interesting since in our placement, the failure of the bottleneck node still results in a stable network topology, whereas if the node nearest to the base-station fails, it is significantly more difficult to recover from.

7. CONCLUSIONS

To summarize, we have formulated an optimization problem that considers jointly node placement, transmission structure and data structure in a data gathering sensor network, in terms of an energy-related cost function. We have studied in detail the 1-D case, namely we provided a closed-form solution for the node placement when the data is independent, and outlined the methodology to solve numerically the case when aggregation at nodes results in either joint entropy coding or Slepian-Wolf coding, for arbitrary correlation structures. We used our insights from the 1-D setting to propose an approximation algorithm that places nodes in a radial "wheel" structure in the 2-D case. We show that significant power gains can be obtained with such a node placement scheme over commonly used uniformly random placements both in terms of total and bottleneck power consumption.

We now briefly mention some caveats of our approach and possible extensions to this work. Our formulation makes some simplifying assumptions regarding radio power consumption that would need to be modified to more accurately capture the characteristics of radios. In practice, radios consume static power for being turned on that would be contribute significantly to the total power consumption for very short range communication. This would add an additional constraint to the formulation presented in Section 3. Our work also assumes that all nodes act both as sensing nodes as well as relays. A possible extension to our work would involve separating the two so that not all relays need to be sensing nodes. Finally, while we constrain the possible topologies to be shortest path trees, the use of other topologies such as the TSP-SPT [Cristescu et al. 2004] might provide additional energy gains in the 2-D case.

REFERENCES

CHENG, P., LIU, X., AND CHUAH, C. N. 2004. Energy-aware node placement in wireless sensor networks. In *IEEE Globecom*.

COVER, T. AND THOMAS, J. 1991. *Elements of Information Theory*. Wiley.

CRESSIE, N. 1991. *Spatial Statistics*. John Wiley and Sons.

CRISTESCU, R., BEFERULL-LOZANO, B., AND VETTERLI, M. 2005. Networked slepian-wolf: Theory, algorithms and scaling laws. IEEE Transactions on Information Theory.

CRISTESCU, R., BEFERULL-LOZANO, B., AND VETTERLI, M. 2004. On network correlated data gathering. In *INFOCOM*.

DASGUPTA, K., KUKREJA, M., AND KALPAKIS, K. 2003. Topology–aware placement and role assignment for energy-efficient information gathering in sensor networks. In *IEEE ISCC*.

EIDENBENZ, S. 2002. Approximation algorithms for terrain guarding. *IPL*.

GANESAN, D., CRISTESCU, R., AND BEFERULL-LOZANO, B. 2004. Power efficient node placement and transmission structure for data gathering under distortion constraints. In *IPSN*.

GOEL, A. AND ESTRIN, D. 2003. Simultaneous optimization for concave costs: single sink aggregation or single source buy-at-bulk. In *ACM-SIAM SODA*.

HAMILTON, M. James San Jacinto Mountains Reserve.

HEINZELMAN, W., CHANDRAKASAN, A., AND BALAKRISHNAN, H. 2000. Energy-efficient communication protocols for wireless microsensor networks. In *HICSS*.

INTANAGONWIWAT, C., GOVINDAN, R., AND ESTRIN, D. 2000. Directed diffusion: A scalable and robust communication paradigm for sensor networks. In *ACM/IEEE Mobicom*. Boston, MA.

LINDSEY, S., RAGHAVENDRA, C., AND SIVALINGAM, K. 2001. Data gathering in sensor networks using the energy*delay metric. In *IPDPS*.

MARCO, D., DUARTE-MELO, E., LIU, M., AND NEUHOFF, D. 2003. On the many-to-one transport capacity of a dense wireless sensor network and the compressibility of its data. In *IPSN*.

POTTIE, G. AND KAISER, W. 2000. Wireless integrated network sensors. *Communications of the ACM 43*, 5 (May), 51–58.

RAPPAPORT, T. S. 1996. *Wireless Communications, Principles and Practice*. Prentice Hall.

SLEPIAN, D. AND WOLF, J. 1973. Noiseless coding of correlated information sources. In *IEEE Trans. Information Theory*. Vol. IT-19. 471–480.

YE, W., HEIDEMANN, J., AND ESTRIN, D. 2002. An energy-efficient mac protocol for wireless sensor networks. In *IEEE Infocom*.