Lifetime Maximization in Wireless Sensor Networks with an Estimation Mission

Iordanis Koutsopoulos
Dept. of Computer and Communications Engineering
University of Thessaly, Greece
Email: jordan@uth.gr

Maria Halkidi
Dept. of Digital Systems
University of Piraeus, Greece
Email: mhalk@unipi.gr

Abstract—We study the problem of maximum lifetime in wireless sensor networks that are entitled with the task of estimating an unknown parameter or process. Sensors take measurements and transfer them in multi-hop fashion to a fusion center (FC) for Maximum Likelihood (ML) estimation. To engineer the network for lifetime maximization while adhering to estimation error specifications, the number of measurements by each sensor per unit of time (namely, sensor measurement rate) and the routes to the FC are controlled. Sensor spatial correlation, measurement accuracies, link qualities and energy reserves affect sensor measurement rates and the routes to the FC, and, in turn, measurement rates and sensor characteristics impact estimation error. We show that the problem can be decomposed into separate optimization problems where each sensor autonomously takes its measurement rate and routing decisions, and we propose an iterative primal-dual algorithm with low-overhead signaling for solving it. Our work optimally captures the fundamental tradeoff between network lifetime and estimation quality and yields a solution based on distributed sensor coordination.

I. Introduction

The primary tasks that are expected to be fulfilled by wireless sensor networks with certain performance guarantees are the estimation of an unknown parameter or parameter process, or the detection of an event in the deployment area. Performance guarantees can be low estimation error, or small probabilities of false alarm and missed detection. Such operational tasks are mapped onto objectives that are different from conventional ones such as end-to-end throughput maximization or delay minimization for which networks are traditionally deployed. Network control actions should take into account these different objectives so that the network performs the tasks it is entitled to with best performance.

We study the interplay between network lifetime and estimation quality. We consider the salient characteristics of wireless sensor networks, most prominent of which is spatial correlation among sensor measurements. Sensor measurement rates need to be controlled, and measurements should be routed to a fusion center (FC) that performs the estimation. The inherent tradeoff is that many measurements improve estimation quality, yet they are inefficient in terms of network lifetime since gathering them to the FC consumes energy. Sensor measurement rate control and routing need to account for spatial correlation of neighboring sensors so that low redundancy data are generated and routed to the FC. Sensor

characteristics, such as energy reserves, measurement quality, and wireless link quality with neighbors shape measurement and routing decisions, so that the rate with which sensor energy is ed is balanced across the network. On the other hand, measurement rates affect estimation quality, since they create different regimes about the joint p.d.f of measurements at the FC and hence affect the estimation error.

A. Related Work

A seminal work in lifetime maximization in wireless sensor networks is [1], where routing of sensor data with certain generation rates to several FCs was addressed. In this work, there is no consideration of estimation constraints or sensor spatial correlation. In information theory, spatially correlated sensors are jointly encoded or compressed to reduce redundant information and transmit minimal necessary amount of bits. In [2], a framework for joint compression of sensor data is proposed for minimizing total cost of data aggregation to sinks. In [3], a similar model is used for the min-cost joint source coding and routing under no contention. In [4], a distributed optimization method is proposed for joint source coding, routing and access control such that total energy is minimized.

Various aspects of estimation objectives are considered [5]-[7]. In [5], the authors study energy minimization through finding the number of quantization levels for given estimation error. The work in [7] studies the impact of transmit power on estimation error. Power allocation for minimum estimation error subject to power constraint, or power minimization subject to an estimation error constraint, possess the water-filling structure. At the optimal solution, only a subset of sensors transmit and the rest are off. Routing and sensor placement to maximize network lifetime was studied in [8] where the distortion depends on sensor positions relative to some cluster-heads. An effort to formulate the lifetime distortion tradeoff with rate - distortion theory was reported in [9]. Finally, [10] studies cooperative routing and defines a link metric to optimally aggregate data for detecting a random field.

B. Our contribution

We study lifetime maximization in multi-hop wireless sensor networks with given estimation error specification. The key new idea is that we control the number of measurements

per unit of time for each sensor, in conjunction with routing to impact estimation error and energy consumption. Our objective can be viewed as compression, but at the measurement packet level, rather than at bit level. By addressing the problem at packet level, we capture the precise dependence between estimation uncertainty due to sensor observation packets and energy efficiency. Our contributions are as follows: (i) we characterize the fundamental tradeoff between estimation accuracy and network lifetime and bring into picture sensor measurement rates as variables, and sensor spatial correlations, (ii) we formulate the problem of maximum lifetime subject to a constraint on estimation error, (iii) we devise an iterative distributed algorithm based on a primal-dual approach, where each sensor separately takes its measurement rate and routing decisions with the aid of feedback from the FC and neighbor sensors in the form of Lagrange multipliers. For the special case of uncorrelated sensor measurements, only feedback from neighbors is needed. En route to the solution, we introduce a metric that captures estimation error for correlated measurements. This is inspired by the total squared cross-correlation (TSC) in code design in CDMA [11]. Our results demonstrate that optimal network lifetime can be achieved through autonomous control of sensor measurement rates and routes based on spatial correlations for given estimation error specification. The rest of the paper is organized as follows. In section II we present the model and assumptions. In section III we formulate the optimization problem and solve it with the distributed algorithm. Section IV presents numerical results and section V concludes our study.

II. SYSTEM MODEL

A. Sensor Measurement and Transmission Model

We consider a sensor network of m sensors, represented by a directed graph $G(\mathcal{N}, \mathcal{A})$, where $\mathcal{N}$ is the set of sensors and $\mathcal{A}$ is the set of links. A link (i, j) between sensors i and j exists if $j \in \mathcal{S}_i$ where $\mathcal{S}_i$ is the set of sensors than can be reached by sensor i with a certain transmit power. Each sensor has initial energy reserve E_i.

We consider a clock-driven system, in which every T_s time units, sensors submit measurements to the FC about a slowly time-varying, unknown, spatially homogeneous phenomenon process. We call each interval of duration T_s an *epoch*. The process under observation is a sequence $\{\theta_t\}_{t=1,2,\ldots}$, where θ_t is the unknown parameter value of the process at epoch t, assumed to remain fixed for the epoch duration. Within a given epoch t, sensor i takes measurements,

$$x_i(\tau) = \theta_\tau + n_i(\tau) \tag{1}$$

where $x_i(\tau)$ is the measurement at time τ and $\theta_\tau = \theta_t$. The noise process $n_i(t)$ captures uncertainty of sensor i observation due to different perception of the process and residual measurement errors. For each i, $n_i(t)$ is Gaussian, zero mean, wide-sense stationary and uncorrelated in time: for any t, t', *temporal correlation* $R_i(t, t') = 0$ if $t \neq t'$ and $R_i(t; t) = \sigma_i^2$ otherwise. The $\sigma_i^2 = \mathbb{E}[n_i^2(t)]$ is the variance

of $n_i(t)$ for any t and captures measurement inaccuracy. Each $n_i(t)$ is independent from $\{\theta_t\}_{t=1,\ldots}$.

Noise processes of each pair of sensors i and j are *spatially correlated* due to sensor proximity. Spatial correlation is time-invariant, namely for any t, t', the spatiotemporal correlation $R_{ij}(s, s'; t, t')$ between sensors i and j at locations s and s' is $R_{ij}(s, s')$ and depends only on locations. By leaving out temporal correlation, we focus on the impact of spatial correlation. Define the symmetric $m \times m$ *spatial correlation matrix* $\mathbf{C}$, whose $[i, j]$-th element, $\rho_{ij} = \mathbb{E}[n_i(t)n_j(t')]$ is the spatial correlation between noise processes $n_i(t)$ and $n_j(t)$ of sensors i and j at all times t, t'. Pairwise correlations ρ_{ij} are assumed non-negative, and they are non-zero only for $j \in \mathcal{S}_i$, and $\mathbf{C}$ is positive definite.

Each epoch consists of a measurement / control interval and a transmission interval. During the former, the FC broadcasts necessary control information to all sensors as detailed in section III. The transmit power of the FC is assumed to be large enough so that all sensors are within its range. During the same interval, each sensor collects measurements. Different measurements of a sensor are uncorrelated in time. However, any measurement of a sensor is *spatially correlated* to any other measurement of another nearby sensor. Also, messaging from neighbor sensors takes place during this phase as detailed in section III. During the transmission interval, sensors forward measurements toward the FC in multi-hop fashion. A sensor forwards portions its data and data it receives from others to its neighbors. We assume no quantization or compression at bit level. The length of the epoch is large enough such that all measurements of sensors reach the FC. At the end of the epoch, the FC makes the estimation.

Let $N_i(t)$ be the number of measurements of sensor i at epoch t. Let the number of epochs T grow, and define $r_i = \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} N_i(t)$ as the *average measurement rate* (in measurements or bits per sec) of sensor i. As discussed later, the vector of measurement numbers $\mathbf{N} = (N_1, \ldots, N_m)$ per epoch or the measurement rate vector $\mathbf{r} = (r_1, \ldots, r_m)$ will be the control variable vector. We refer to both terms interchangeably in the paper and consider them continuous. We assume that processes $\{N_i(t)\}$, $\{N_j(t)\}$ for $i \neq j$ are uncorrelated in time.

A sensor consumes energy only during transmission. The energy consumed by i to transmit an information unit (e.g. a measurement packet) to sensor $j \in \mathcal{S}_i$ is e_{ij}. This captures wireless link gain between i and j, namely path loss and fading and well as the requirement for a sufficient signal to noise ratio (SNR) at a receiver. Let e_i denote the energy consumed per information unit by sensor i to the FC.

B. Estimation at the FC

Within each epoch, the FC obtains sensor measurements and computes an estimate $\hat{\theta}_t$ of θ_t in the Maximum Likelihood (ML) sense. ML is a valid estimation in the absence of prior knowledge about θ_τ. We drop epoch index t. First, let each sensor send one measurement to the FC, so that an ensemble of m measurements $\{x_i\}_{i=1}^{m}$ is available. The joint

p.d.f. of sensor measurements $\mathbf{x} = (x_1, \ldots, x_m)^T$, $p_\theta(\mathbf{x})$, is Gaussian with mean $\theta\mathbf{1}$ and correlation matrix $\mathbf{C}$, where $\mathbf{1}$ is the $m \times 1$ vector of all ones. The ML estimate of θ, is $\hat{\theta}_{ML} = \arg\max_\theta \log p_\theta(\mathbf{x}) = (\mathbf{1}^T\mathbf{C}^{-1}\mathbf{x})/(\mathbf{1}^T\mathbf{C}^{-1}\mathbf{1})$. The criterion for estimation quality is mean squared error (MSE), $\mathbb{E}[(\theta - \hat{\theta}_{ML})^2]$. To make estimation quality independent of θ, we consider unbiased estimators, i.e., $\mathbb{E}(\hat{\theta}_{ML}) = \theta$. Then, MSE equals $\mathrm{var}(\hat{\theta}_{ML})$. We seek to minimize $\mathrm{var}(\hat{\theta}_{ML})$ (where expectation is with respect to randomness of observations), through a minimum variance unbiased (MVU) estimator. In our case, the ML estimate is unbiased, and $\mathrm{var}(\hat{\theta}_{ML}) = [\mathbf{1}^T\mathbf{C}^{-1}\mathbf{1}]^{-1}$. For spatially uncorrelated sensors, $\hat{\theta}_{ML} = \left(\sum_{i=1}^m \frac{1}{\sigma_i^2}\right)^{-1} \cdot \left(\sum_{i=1}^m \frac{x_i}{\sigma_i^2}\right)$ and $\mathrm{var}(\hat{\theta}_{ML}) = \left(\sum_{i=1}^m \frac{1}{\sigma_i^2}\right)^{-1}$.

III. MAX-LIFETIME ROUTING AND MEASUREMENT RATE CONTROL SUBJECT TO ESTIMATION ERROR CONSTRAINTS

A. Estimation Error vs. Sensor Numbers of Measurements

We first obtain the joint p.d.f. of sensor measurements and then proceed to the expression of $\mathrm{var}(\hat{\theta}_{ML})$ for different numbers of measurements per sensor.

1) Spatially Uncorrelated Sensors: Assume that at a given epoch, sensor i gets N_i measurements, $i = 1, \ldots, m$ collected in vector $\mathbf{x}^{(i)} = (x_1^{(i)}, \ldots, x_{N_i}^{(i)})$. Let n_i, $i = 1, \ldots, m$ be indices with $1 \le n_i \le N_i$. Measurements of each sensor are uncorrelated among themselves. Since sensors are spatially uncorrelated, all sensor measurements are uncorrelated, and under the Gaussian assumption, they are independent. The ML estimate of θ, $\hat{\theta}_{ML}$, is

$$\hat{\theta}_{ML} = \left(\sum_{i=1}^m \sum_{n_i=1}^{N_i} \frac{x_{n_i}^{(i)}}{\sigma_i^2}\right) \cdot \left(\sum_{i=1}^m \frac{N_i}{\sigma_i^2}\right)^{-1}, \quad (2)$$

and $\mathrm{var}(\hat{\theta}_{ML}) = \left(\sum_{i=1}^m \frac{N_i}{\sigma_i^2}\right)^{-1}$. Thus, estimation error depends on the number of measurements N_i of each sensor i.

2) Spatially Correlated Sensors: The situation here is more complex, due to correlations among measurements of different sensors. We start from the simple case of two sensors, 1 and 2, with $N_1 = 1$ and $N_2 = 2$. The joint p.d.f. of vector $\mathbf{x} = (x_1^{(1)}, x_1^{(2)}, x_2^{(2)})$ can be found by arguing as follows:

$$p_\theta(x_1^{(1)}, x_1^{(2)}, x_2^{(2)}) = p_\theta(x_1^{(2)}|x_1^{(1)}) p_\theta(x_2^{(2)}|x_1^{(1)}) p_\theta(x_1^{(1)}), \quad (3)$$

since measurements of sensor 2, $x_1^{(2)}, x_2^{(2)}$ are conditionally independent, given the measurement $x_1^{(1)}$ of 1. Finally,

$$p_\theta(x_1^{(1)}, x_1^{(2)}, x_2^{(2)}) = \frac{p_\theta(x_1^{(1)}, x_1^{(2)}) p_\theta(x_1^{(1)}, x_2^{(2)})}{p_\theta(x_1^{(1)})}, \quad (4)$$

and the joint p.d.f. $p_\theta(x_1^{(1)}, x_1^{(2)}, x_2^{(2)})$ can be expressed in terms of distributions $p_\theta(x_1^{(1)}, x_1^{(2)})$ and $p_\theta(x_1^{(1)}, x_2^{(2)})$ which are jointly Gaussian with correlation matrix equal to the 2×2 spatial correlation matrix of sensors 1 and 2, while $p_\theta(x_1^{(1)})$ is Gaussian with variance σ_1^2.

Similarly, the joint p.d.f. of measurement vectors $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}$ where sensors take $N_1, N_2 > 1$ measurements, is

$$p_\theta(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) = \frac{\prod_{n_1=1}^{N_1} \prod_{n_2=1}^{N_2} p_\theta(x_{n_1}^{(1)}, x_{n_2}^{(2)})}{\left(\prod_{n_1=1}^{N_1} p_\theta(x_{n_1}^{(1)})\right)^{N_2-1} \left(\prod_{n_2=1}^{N_2} p_\theta(x_{n_2}^{(2)})\right)^{N_1-1}}. \quad (5)$$

and the variance of estimation error is

$$\mathrm{var}(\hat{\theta}_{ML}) = \left(N_1 N_2 \mathbf{1}^T\mathbf{C}^{-1}\mathbf{1} - \frac{(N_2-1)N_1}{\sigma_1^2} - \frac{(N_1-1)N_2}{\sigma_2^2}\right)^{-1}, \quad (6)$$

where $\mathbf{1}$ is the 2×1 vector of 1's.

Consider the general case of m sensors, where sensor i takes N_i measurements. For sensor subset $\mathcal{S} \subseteq \mathcal{M}$, define as $\mathbf{C}_\mathcal{S}$ the sub-matrix of $\mathbf{C}$ that consists only of rows and columns corresponding to sensors in $\mathcal{S}$. Let $|\mathcal{S}|$ be the cardinality of $\mathcal{S}$. After some tedious algebra, we obtain

$$\mathrm{var}(\hat{\theta}_{ML}) = \left[\sum_{\mathcal{S} \subseteq \mathcal{M}} \gamma_\mathcal{S} \left(\prod_{i \in \mathcal{S}} N_i \prod_{j \in \mathcal{M} \setminus \mathcal{S}} (N_j - 1)\right) \mathbf{1}^T\mathbf{C}_\mathcal{S}^{-1}\mathbf{1}\right]^{-1} \quad (7)$$

where $\gamma_\mathcal{S} = (-1)^{|\mathcal{S}|}$, if the number of sensors m is even, and $\gamma_\mathcal{S} = (-1)^{|\mathcal{S}|+1}$ if m is odd, and $\mathbf{1}$ is a vector of ones of appropriate dimension.

Let $\mathrm{var}(\hat{\theta}_{ML}) = [h(\mathbf{N})]^{-1}$. Each term in the sum in $h(\mathbf{N})$ denotes mutual coupling among measurements of a subset $\mathcal{S}$ of sensors. To compute $\mathrm{var}(\hat{\theta}_{ML})$, one needs to compute $2^m - 1$ terms, each of which involves finding $C_\mathcal{S}^{-1}$. Depending on the deployment application, the number of sensors vary from less than ten to some hundreds. For few sensors, $\mathrm{var}(\hat{\theta}_{ML})$ can be precisely evaluated. For larger number of sensors, some methods are needed to reduce computation load. We could split the set of sensors $\mathcal{M}$ into subsets by clustering, compute precisely an estimate and error variance for each sensor subset, and fuse all estimates.

We introduce a solution inspired from CDMA code design. Each user code is a vector, and the pairwise code cross-correlation is their inner product. The total squared cross-correlation (TSC) metric, defined as the sum of squares of pairwise code cross-correlations quantifies mutual interference among CDMA codes and is used in designing low-interference, high capacity systems [11]. From (6), note that for two sensors i and j, it is $\mathrm{var}(\hat{\theta}_{ML}) = \left(N_i N_j \alpha_{ij} + \frac{N_i}{\sigma_i^2} + \frac{N_j}{\sigma_j^2}\right)^{-1}$, with $\alpha_{ij} = \mathbf{1}^T\mathbf{C}_{ij}^{-1}\mathbf{1} - \frac{1}{\sigma_i^2} - \frac{1}{\sigma_j^2}$, and $\mathbf{C}_{ij}$ the 2×2 correlation matrix that corresponds to sensors i and j. Thus, α_{ij} may be viewed as the coupling between measurements of sensors i and j; if $\rho_{ij} = 0$, then $\alpha_{ij} = 0$. We construct the following metric for estimation error variance, which we call *Total Pairwise Correlation Approximation (TPCA)* of variance of estimation error,

$$\mathrm{TPCA}(\hat{\theta}_{ML}) = \left(\sum_{i=1}^m \sum_{j \in \mathcal{S}_i, j>i}^m N_i N_j \alpha_{ij} + \sum_{i=1}^m \frac{N_i}{\sigma_i^2}\right)^{-1} = \tilde{h}(\mathbf{N})^{-1}. \quad (8)$$

B. Problem Statement and Formulation

The fundamental tradeoff between estimation quality and lifetime is that more measurements yield small estimation error but require more energy to gather at the FC. Since sensors have different observation quality, energy reserves and consumed energy per measurement, it is necessary to provide sensors with the possibility to submit *different* average number of measurements per unit time, i.e different measurement rates. When sensor spatial correlations come into play, sensor measurement rates should be controlled, so that sensors in closer proximity collectively transmit fewer measurements than if sensors were uncorrelated to avoid data redundancy. Due to the coupling among measurements of different sensors because of spatial correlation, different measurement rate vectors give rise to different joint probability distribution of measurements, and different estimation errors. By controlling measurement rates, we create sensor transmission regimes that provide the FC with adequate measurements for the specified estimation error while adhering to low energy consumption. Routing amounts to selecting the portions of measurements to transmit to each neighbor. Each sensor forwards its own data as well as received data from other sensors. The joint problem is to determine the measurement rate of each sensor (which essentially denotes injected traffic rate in the network) and the routes through which measurements are transferred from sensors to the FC such that network lifetime is maximized.

We adopt a fluid model for information flow, so that information forwarded between sensors i and $j \in \mathcal{S}_i$ is a real-valued flow $f_{ij} \geq 0$ denoting average amount of bits per unit of time. Let $\mathbf{f} = (f_{ij} : (i,j) \in \mathcal{A})$ be the vector of all link flows. The lifetime of sensor i is $L_i(\mathbf{f}) = E_i/(\sum_{j \in \mathcal{S}_i} e_{ij} f_{ij})$, where the denominator denotes energy consumption rate of sensor i. Network lifetime is defined as the time until the battery of the first sensor empties, namely it is $\min_{i \in \mathcal{N}} L_i(\mathbf{f})$. Besides, $\mathbf{f}$, the measurement rate vector $\mathbf{r}$ of sensors needs to be determined.

We are given a constraint ε on average estimation error. Expression $[\tilde{h}(\mathbf{N}^{(t)})]^{-1}$ in (8) gives the estimation error within an epoch as function of numbers of measurements. To obtain an expression for the average estimation error over all epochs, $\bar{h}(\mathbf{r}) = \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} \tilde{h}(\mathbf{N}^{(t)})$, we need to express it through $\mathbf{r}$. It can be shown that,

$$\bar{h}(\mathbf{r}) = T_s^2 \sum_i \sum_{j \in \mathcal{S}_i, j > i} \alpha_{ij} r_i r_j + T_s \sum_i \frac{r_i}{\sigma_i^2}. \quad (9)$$

The problem can be formulated as:

$$\max_{\mathbf{r}, \mathbf{f}} \min_{i \in \mathcal{N}} L_i(\mathbf{f})$$

subject to the the flow conservation constraint at each sensor, and the estimation error constraint,

$$\sum_{j:i \in \mathcal{S}_j} f_{ji} + r_i = \sum_{j \in \mathcal{S}_i} f_{ij}, \quad \forall i \in \mathcal{N}$$

$$\frac{T_s^2}{2} \sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{S}_i} \alpha_{ij} r_i r_j + T_s \sum_{i \in \mathcal{N}} \frac{r_i}{\sigma_i^2} = \frac{1}{\varepsilon},$$

with $\mathbf{r} \geq \mathbf{0}$, $\mathbf{f} \geq \mathbf{0}$. Note that we multiply the last constraint by $1/2$ to transform it into a form that does not need additional coordination among sensors. Define a new variable $z = \min_{i \in \mathcal{N}} L_i(\mathbf{f})$, and set $\hat{\mathbf{r}} = z\mathbf{r}$ and $\hat{\mathbf{f}} = z\mathbf{f}$ to get the equivalent formulation,

$$\max_{\hat{\mathbf{r}}, \hat{\mathbf{f}}, z} z$$

$$\text{subject to: } \sum_{j:i \in \mathcal{S}_j} \hat{f}_{ji} + \hat{r}_i = \sum_{j \in \mathcal{S}_i} \hat{f}_{ij}, \quad \forall i \in \mathcal{N},$$

$$\frac{T_s^2}{2} \sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{S}_i} \hat{r}_i \hat{r}_j \alpha_{ij} + T_s \sum_{i \in \mathcal{N}} \frac{\hat{r}_i z}{\sigma_i^2} = \frac{z^2}{\varepsilon} \quad (10)$$

$$\sum_{j \in \mathcal{S}_i} e_{ij} \hat{f}_{ij} \leq E_i, \quad \forall i,$$

and $\hat{\mathbf{r}} \geq \mathbf{0}$, $\hat{\mathbf{f}} \geq \mathbf{0}, z > 0$, where the last constraint is due to the new variable z.

C. Primal-Dual Algorithm

We relax the flow conservation and estimation error constraints and form the Lagrangian

$$L(z, \hat{\mathbf{r}}, \hat{\mathbf{f}}, \boldsymbol{\lambda}, \mu) = -z + \sum_{i=1}^{m} \lambda_i \left(\sum_{j:i \in \mathcal{S}_j} \hat{f}_{ji} + \hat{r}_i - \sum_{j \in \mathcal{S}_i} \hat{f}_{ij} \right)$$

$$+ \mu \left(\frac{T_s^2}{2} \sum_{i=1}^{m} \sum_{j \in \mathcal{S}_i} \hat{r}_i \hat{r}_j \alpha_{ij} + T_s \sum_{i=1}^{m} \frac{\hat{r}_i z}{\sigma_i^2} - \frac{z^2}{\varepsilon} \right)$$

where $\boldsymbol{\lambda}$, μ are the Lagrange multipliers for flow conservation constraints and the estimation error constraint. The primal problem,

$$\min_{z > 0, \hat{\mathbf{r}} \geq \mathbf{0}, \hat{\mathbf{f}} \geq \mathbf{0}} L(z, \hat{\mathbf{r}}, \hat{\mathbf{f}}, \boldsymbol{\lambda}, \mu) \quad \text{s.t.} \quad \sum_{j \in \mathcal{S}_i} e_{ij} \hat{f}_{ij} \leq E_i, \quad \forall i,$$
$$(11)$$

is divided into a routing problem that depends on $\hat{\mathbf{f}}$, and a part is a function of $\hat{\mathbf{r}}, z$, call it $g(\hat{\mathbf{r}}, z)$. The routing problem decomposes into linear programming (LP) problems that each sensor i solves to determine next-hop forwarding variables $\hat{f}_{ij}$,

$$\min_{\hat{f}_{ij}} \sum_{j \in \mathcal{S}_i} (\lambda_j - \lambda_i) \hat{f}_{ij} \quad \text{s.t.} \quad \sum_{j \in \mathcal{S}_i} e_{ij} \hat{f}_{ij} \leq E_i. \quad (12)$$

Note that for (12), Lagrange multipliers from sensors $j \in \mathcal{S}_i$ need to be communicated to i. The second problem is to minimize $g(\hat{\mathbf{r}}, z)$. The optimization of $g(\hat{\mathbf{r}}, z)$ with respect to z is carried by the FC by setting $\partial g(\hat{\mathbf{r}}, z)/\partial z = 0$ to find $z^{(t)}$. Next, each sensor adjusts $\hat{r}_i$ by running one iteration of the gradient algorithm towards minimizing $g(\hat{\mathbf{r}}, z)$, that is,

$$\hat{r}_i^{(t)} = \hat{r}_i^{(t-1)} - s_t \partial g(\hat{\mathbf{r}}^{(t-1)}, z^{(t)})/\partial \hat{r}_i, \quad (13)$$

where $s_t > 0$ is the step size. For this iteration, a sensor i needs to know instantaneous $\hat{r}_j$ of neighbor sensors.

The dual problem is:

$$\max_{\boldsymbol{\lambda}, \mu} L(z, \hat{\mathbf{r}}, \hat{\mathbf{f}}, \boldsymbol{\lambda}, \mu) \quad (14)$$

for given $\hat{\mathbf{f}}, z, \hat{\mathbf{r}}$. Sensor i knows instantaneous incoming and outgoing flow variables $\hat{f}_{ji}$ and $\hat{f}_{ij}$ from and to neighboring sensors, and updates its Lagrange multiplier λ_i according to:

$$\lambda_i^{(t)} = \lambda_i^{(t-1)} + s_t \Big(\sum_{j:i\in\mathcal{S}_j} \hat{f}_{ji}^{(t-1)} + \hat{r}_i^{(t-1)} - \sum_{j\in\mathcal{S}_i} \hat{f}_{ij}^{(t-1)} \Big) \quad (15)$$

Next, for given $\hat{\mathbf{r}}$, the FC updates Lagrange multiplier μ,

$$
\begin{aligned}
\mu^{(t)} &= \mu^{(t-1)} + s_t \Big(\frac{T_s^2}{2} \sum_{i=1}^{m} \sum_{j\in\mathcal{S}_i} \hat{r}_i^{(t-1)} \hat{r}_j^{(t-1)} \alpha_{ij} \\
&+ T_s \sum_{i=1}^{m} \frac{\hat{r}_i^{(t)} z^{(t-1)}}{\sigma_i^2} - \frac{z^{(t-1)2}}{\varepsilon} \Big).
\end{aligned}
\quad (16)
$$

The primal-dual algorithm, is summarized below.

- **STEP 0:** Initialization. Before epoch $t = 0$, FC initializes $\mu^{(0)}$ and $z^{(0)} > 0$. Each sensor i initializes $\hat{r}_i^{(0)}$ and $\lambda_i^{(0)}$.
- **IF $t > 0$ DO**
- **STEP 1:** Each sensor i updates $\lambda_i^{(t)}$ according to (15).
- **STEP 2:** The FC updates Lagrange multiplier, $\mu^{(t)}$ as in (16).
- **STEP 3:** The FC computes the new variable, $z^{(t)}$ from $\partial g(\hat{\mathbf{r}}, z)/\partial z = 0$ as

$$z^{(t)} = \frac{\varepsilon}{2} \Big(T_s \sum_{i=1}^{m} \frac{\hat{r}_i^{(t-1)}}{\sigma_i^2} - \frac{1}{\mu^{(t)}} \Big). \quad (17)$$

- **STEP 4:** The FC broadcasts $z^{(t)}$ and $\mu^{(t)}$ to all sensors.
- **STEP 5:** Each sensor i updates its variable $\hat{r}_i$ with the gradient iteration

$$\hat{r}_i^{(t)} = \hat{r}_i^{(t-1)} - s_t \Big[\lambda_i^{(t)} + \mu^{(t)} \Big(\frac{T_s^2}{2} \sum_{j\in\mathcal{S}_i} \alpha_{ij} \hat{r}_j^{(t-1)} + \frac{T_s z^{(t)}}{\sigma_i^2} \Big) \Big] \quad (18)$$

if $\hat{r}_i^{(t)} > 0$, else $\hat{r}_i^{(t)}$ is set to 0.
- **END IF**
- **STEP 6:** Each sensor i sets its generation rate as $r_i^{(t)} = \hat{r}_i^{(t)}/z^{(t)}$, or equivalently it generates $N_i^{(t)} = T_s r_i^{(t)}$ measurements during epoch t. Then it broadcasts $r_i^{(t)}$ to its neighbors.
- **STEP 7:** Each sensor i solves the routing problem (12) and computes $\hat{f}_{ij}^{(t)}$, $j \in \mathcal{S}_i$. It determines its flows as $f_{ij}^{(t)} = \hat{f}_{ij}^{(t)}/z^{(t)}$.
- **STEP 8:** Measurements from sensors reach the FC during epoch t.
- **STEP 9** $t \leftarrow t + 1$. Go to Step 1. Continue until convergence.

The algorithm is decentralized, with minimal feedback from the FC. At each epoch, dual variables $\boldsymbol{\lambda}^{(t)}$ and $\mu^{(t)}$ are updated: each sensor computes $\lambda_i^{(t)}$ and the FC updates $\mu^{(t)}$. These are performed each with a gradient ascent step. Next, a one-shot minimization is performed by the FC to find $z^{(t)}$. Then, each sensor adjusts $\hat{r}_i^{(t)}$ related to generation rate with a gradient descent step after obtaining $\hat{r}_j$ from neighbors. In Step 5, each sensor essentially calibrates local measurement generation rates based on spatial correlation with neighbors.
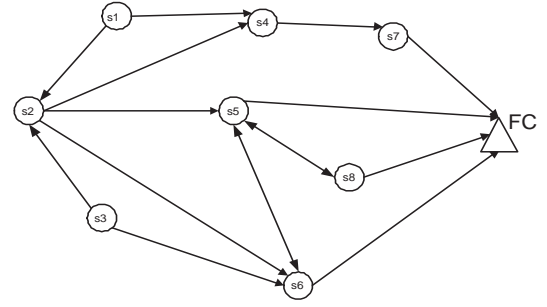

Fig. 1. Topology of a network with 8 sensors.The distance between two neighboring sensors varies randomly between 1 and 8m.

Finally, each sensor independently takes routing decisions by solving the LP problem (12). If the sequence of steps satisfies $\lim_{t\to\infty} s_t = 0$ and $\sum_t s_t = \infty$, the algorithm converges to at least a local optimum of the original problem. The problem structure is reminiscent of the classical rate control and routing problem for network utility maximization in wireless networks [12], albeit with a different objective and the non-conventional constraint of estimation error.

1) Uncorrelated sensors: For the special case of loosely correlated or uncorrelated sensors, the constraint on estimation error becomes

$$T_s \sum_i \frac{\hat{r}_i}{\sigma_i^2} = \frac{z}{\varepsilon}. \quad (19)$$

It can be shown that the problem decomposes into an LP problem, that is further decomposed into a measurement rate update for each sensor i through the iteration

$$\hat{r}_i^{(t)} = \Big[\hat{r}_i^{(t-1)} + \Big(\lambda_i^{(t)} - \frac{1}{\sigma_i^2} \Big) \Big]^+$$

(where $x^+ = x$ if $x > 0$, else it is 0), and a routing problem as in (12). In this case, the primal-dual algorithm can be shown to converge to the global optimum routing and measurement generation rates.

IV. NUMERICAL RESULTS

We consider a network of $m = 8$ sensors (see Figure 1), where the distance between two neighboring sensors varies randomly between 1 and 8m. For spatial correlation, we adopt the power exponential model [13] according to which two sensors at distance d meters have $\rho(d) = \exp[-(\frac{d}{K_1})^{K_2}]$, with $K_1 \in [2, 6]$ and $K_2 = 1$. Then, the correlation of sensors at a specific distance (d) increases as we increase the value of $K1$. We adopt the transmission model from [14], where the energy consumed to reliably transmit a packet to the FC at distance d is $100\text{nJ} \times d^2$. Thus, spatial correlation between neighbor sensors varies from about 0.2 to 0.8 and e_{ij} varies from 0.1 to $6.4\mu\text{J}$. Non-neighboring sensors are uncorrelated. Each sensor has initial energy reserve $E = 1$ Joule, $\sigma^2 = 1$, and transmits packets of 100 bits. For each sensor, we consider only the energy consumed for transmission. For the estimation error, we use the expression (8). Simulation results were generated in Matlab. In Figure 2 (a) and (b), we plot system lifetime as function of K_1 for error tolerance ϵ equals to 10^{-3} and

10^{-5}. We average results over 50 experiments for various randomly defined distances between neighboring sensors while the value of K_1 varies from 2 to 6. Lifetime increases as K_1 increases, that is as sensors become more correlated. Each pair of sensors in the network jointly submits fewer measurements than the total number of measurements per pair if the sensors were uncorrelated. This implies that the flow rates between sensors are reduced which also results in the reduction of energy consumption in the network and therefore the network lifetime is prolonged. We can also say that for same spatial correlation, the lifetime for looser constraints on estimation error ($\varepsilon = 10^{-3}$) is at least one order of magnitude more than that for $\varepsilon = 10^{-5}$. Also, as the correlation of sensor increases (i.e. K_1 increases) this difference becomes more clear. Moreover we observe that for strict constraints on error tolerance (Figure 2 (b)), the lifetime does not significantly vary with correlation. This makes sense since the effect of a strict constraint, such as $\varepsilon = 10^{-5}$, to the network lifetime overcovers that of correlation. Due to low error tolerance, the sensors need to send a large number of measurements to the FC in order that a more accurate estimation is achieved. The transmission of measurements to the FC consumes energy which significantly affects the system lifetime.

V. DISCUSSION

We addressed the problem of maximizing lifetime in a sensor network subject to a constraint on the estimation error. Our main contribution is the joint orchestration of sensor measurement rates and routes to the FC, and the design of a distributed primal-dual algorithm that relies on lightweight local sensor coordination and feedback from the FC. Sensor spatial correlation is used so that sensors control their measurement rates to avoid redundant data generation. This work adhered to the simple scenario where sensors generate and forward their and others' measurements with no further processing, in order to better demonstrate the way measurement rates and routes affect estimation quality and lifetime. In that sense, it can be considered as a prelude to a more general approach that will include in-network aggregation by sensors prior to transmission in an effort to better shape the tradeoff between estimation quality and energy efficiency. We plan to address this issue in a future work.

(a)

(b)

Fig. 2. System Lifetime versus the parameter K_1 used in the definition of sensor correlation ($\rho(d) = \exp[-(\frac{d}{K_1})^{K_2}]$) for error tolerance a) $\epsilon = 10^{-3}$ and b) $\epsilon = 10^{-5}$.

REFERENCES

[1] J.-H. Chang and L. Tassiulas, "Maximum lifetime routing in wireless sensor networks", *IEEE/ACM Trans. Networking,* vol.12, no.4, pp.609-619, Aug. 2004.

[2] S.J. Baek, G. de Veciana and X. Su, "Minimizing energy consumption in large-scale sensor networks through distributed data compression and hierarchical aggregation", *IEEE J. Select. Areas Commun.,* vol.22, no.6, pp.1130-1140, Aug. 2004.

[3] R. Cristescu, B. Beferull-Lozano, M. Vetterli and R. Wattenhofer, "Network correlated data gathering with explicit communication: NP-Completeness and algorithms", *IEEE/ACM Trans. Networking,* vol.14, no.1, pp.41-54, 2006.
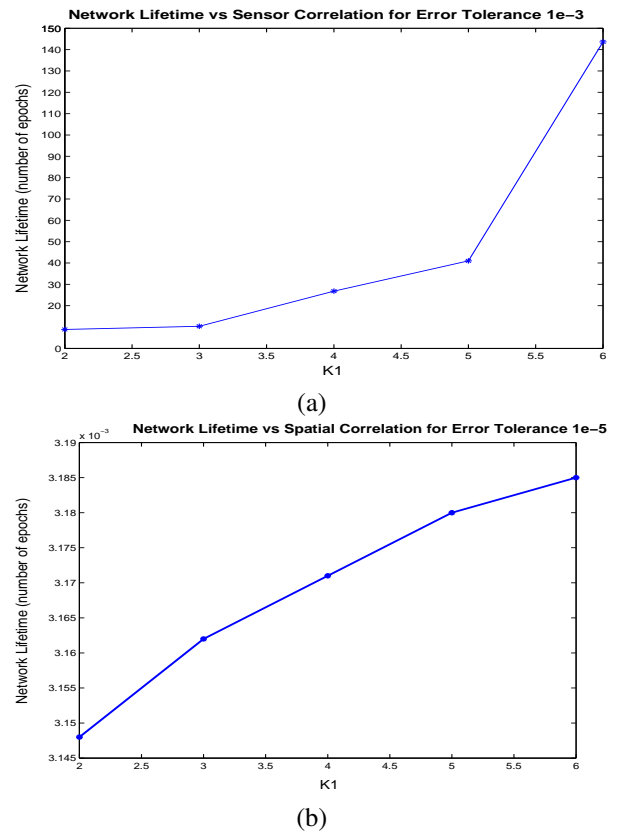
[4] K. Yuen, B. Liang and B. Li, "A distributed framework for correlated data gathering in sensor networks", *IEEE Trans. Veh. Tech.,* vol.57, no.1, pp.578-593, 2008.

[5] A. Krasnopeev, J.-J. Xiao and Z.-Q. Luo, "Minimum energy decentralized estimation in a wireless sensor network with correlated sensor noise", *EURASIP J. Wireless Com. Networking,* vol.5, no.4, pp.473-482, 2005.

[6] J. Matamoros and C. Anton-Haro, "Opportunistic random access for distributed parameter estimation in wireless sensor networks", *Europ. Signal Proc. Conf. (EUSIPCO),* Poznan, Poland 2007.

[7] S. Cui, J. Xiao, A. Goldsmith, Z. Luo and H. V. Poor, "Estimation Diversity and Energy Efficiency in Distributed Sensing", *IEEE Trans. Signal Proc.,* vol.55, no.9, pp.4683-4695, Sept. 2007.

[8] V. Sharma, E. Frazzoli and P.G. Voulgaris, "Improving lifetime data gathering and distortion for mobile sensing networks", *Proc. IEEE SECON,* 2004.

[9] A. Kansal, A. Ramamoorthy, M.B. Srivastava and G.J. Pottie, "On sensor network lifetime and data distortion", *Proc. IEEE Int. Symp. on ISIT,* 2005.

[10] Y. Sung, S. Misra, L. Tong and A. Ephremides, "Cooperative routing for distributed detection in large sensor networks", *IEEE J. Select. Areas Commun., vol.25, no.2, pp.471-483,* Feb. 2007.

[11] J.L. Massey and T. Mittelholzer, "Welch's bound and sequence sets for code-division multiple access systems", in *Sequences II, Methods in Communication, Security and Computer Science,* A. Capocelli, A. De Santis and U. Vaccaro, Eds. New York: Springer-Verlag, 1993.

[12] X. Lin, N.B. Shroff and R. Srikant, "A tutorial on cross-layer optimization in wireless networks", *IEEE J. Select. Areas Commun, vol.24, no.8, pp.1452-1463,* Aug. 2006.

[13] M.C. Vuran, O.B. Akan and I.F. Akyildiz, "Spatio-Temporal correlation: Theory and applications for wireless sensor networks," *Computer Net. Journal (Elsevier),* vol.45, no.3, pp. 245-261, June 2004.

[14] W. Heinzelman, A.P. Chandrakasan, and H. Balakrishnan, "Energy-Efficient Communication Protocols for Wireless Microsensor Networks", *Proc. HISS,* 2000.