ADAPTIVE QUANTIZATION FOR MULTIHOP PROGRESSIVE ESTIMATION IN WIRELESS SENSOR NETWORKS

Santosh Shah, Baltasar Beferull-Lozano

GSIC, Instituto de Robótica y Tecnologías de la Información y las Comunicaciones
Universidad de Valencia, 46980, Paterna (Valencia), Spain
{Santosh.Shah, Baltasar.Beferull}@uv.es

ABSTRACT

We consider the problem of parameter estimation in a wireless sensor network, where because of the bandwidth and power constraints, each sensor transmits quantized information to its parent on a multihop path. Our approach jointly optimizes: i) sensor selection, ii) routing structure, and iii) number of bits per sample for each sensor. First, we express our problem as an optimization problem, and then we design an algorithm that is based on an adaptive quantization and an estimate-and-forward scheme that allows performing sequentially this joint optimization in an efficient way. We show that our algorithm provides a routing structure that trades-off the aforementioned three metrics better than the traditional shortest path tree routing structure, which is commonly used in practice. Numerical results depicting the performance and advantages of our approach over previous state-of-the-art approaches are presented.

1. INTRODUCTION

The power efficiency in a Wireless Sensor Network (WSN) can be achieved by adopting multihop routing and by activating only a subset of sensors at a desired time slot. Then, each selected sensor should apply an Estimate-and-Forward (EF) scheme, presented in [1, 2], to fuse all other measurements that are received from its child sensors together with its own measurement to perform the parameter estimation, and then transmit only one flow of data to its parent sensor in the chosen routing structure. Given a WSN with a coverage graph and a sink (querying) node, we consider a problem in which three metrics: i) total distortion in estimation, ii) total communication cost, and iii) number of bits per sample for each sensor are optimized jointly. Thus, for a given total power budget the best subset of sensors, the optimal bit allocation, and the best associated routing structure to send the aggregated information to the sink node is achieved. Furthermore, when considering the EF scheme and allowing an interplay among these three metrics, we will show an important result that the routing structure using the traditional Shortest Path Tree (SPT) [3], widely used in practice, is no longer optimal. That is, our solution for routing provides a better trade-off among these three metrics.

We consider another scheme, which we will take into account to compare the performance of our work, is an Measure-and-Forward (MF) [4]. In this scheme, measurements and quantization errors are simply forwarded via a multihop path to the sink node for final estimation. The authors in [4], consider the network lifetime maximization issue for an estimation application in energy-limited sensor networks, where a Linear Programming (LP) problem is formulated to minimize jointly the total number of bits at each sensor and the total transmission cost of the multihop routing to the sink node. Since it applies the MF scheme, we name it as an LP-MF. Notice that the total throughput generated in the MF scheme is always larger than compared to the EF scheme, and therefore the total transmission cost in the MF scheme is always larger. In [5], a Progressive Distributed Estimation using the EF scheme is proposed, we name it as a PDE-EF, where the objective is to estimate an unknown deterministic parameter using the Best Linear Unbiased Estimator (BLUE) estimator [6] in order to save the total energy. The main demerit of this approach is that there is no joint trade-off among all three metrics and it does not provide the choice of selecting a subset of sensors.

We compare our approach with the previously presented state-of-the-art approaches: LP-MF [4] and PDE-EF [5], showing a better performance. The organization of the paper is as follows: Section II presents our problem formulation for progressive estimation following with our optimization problem. Section III describes our adaptive quantization based approaches. We present our experimental results in section IV. Finally, conclusions are given in section V.

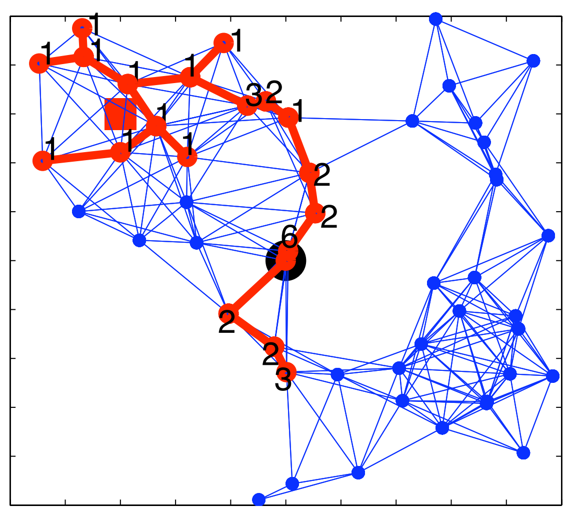

Fig. 1: A WSN simulation example with $N = 50$ sensors where the following elements can be seen: a subset of selected sensors; an associated routing structure (thick edges); and number of bits per sample to quantize measurements of the selected sensors (numbers in selected sensors), and where a sink node and a source target are located at the center (large black circle) and at the top left corner (square), respectively. Thin edges represent the underlying network connectivity graph and thick edges belong to the selected routing structure; larger communication cost links, for which $d_{j,k} > d_{th}$ are eliminated, where $d_{j,k}$ is the distance between sensor j and k.

This work was supported by the Spanish MEC Grants TEC2010-19545-C04-04 "COSIMA", CONSOLIDER-INGENIO 2010 CSD2008-00010, "COMONSENS", the European STREP Project "HYDROBIONETS" Grant no. 287613 within the FP7 Framework Programme, and by a Telefonica Chair.

2. PROBLEM FORMULATION

We consider a WSN consisting of n sensors with a unique identity $j \in \{1, 2, ..., n\}$ and there is a sink node with an identity $n+1$, where final estimation is to be performed under given total power constraint. Further, we consider that sensors are deployed in a square region and the distance between a sensor and any potential neighbor can be expressed as ξd_{th}, where $\xi \in (0, 1]$ is uniformly distributed and d_{th} is a normalizing factor. This results in a network connectivity symmetric graph *G = (V, E)*, as illustrated in Fig. 1, where *V* is a set of n stationary sensors and *E* is a set of communication links. We assume that there is an underlying MAC protocol available, which resolves the network collisions. We also assume that an incidence matrix A is given such that: $A_{i,j} = 1$ if $d_{i,j} \leq d_{th}$ that implies that the links (i, j) and (j, i) exist in the network connectivity graph and $A_{i,j} = 0$ otherwise, where $d_{i,j}$ is the distance between sensor i and sensor j. Let us assume that sensor j makes an observation $y_j \in \mathbb{R}$ on a deterministic parameter $\theta \in \mathbb{R}$, and is described by:

$$y_j = h_j\theta + z_j, \quad j = 1, 2, \ldots, n \tag{1}$$

where θ is the unknown deterministic scalar parameter to be estimated, whose observation is distorted by a scalar $h_j \in \mathbb{R}$ and corrupted by additive noise z_j, which is assumed to be i.i.d. with pdf $\mathcal{N}(0, \sigma^2_{z_j})$, and where $\sigma^2_{z_j}$ is assumed to be spatially varying, but otherwise unknown. We consider that the scalar h_j follows the form $h_j = d_{j,t}^{-\beta}$, that is, $y_j = \theta/d_{j,t}^{\beta} + z_j$. Here, h_j can be a signal strength decay model [7], where $d_{j,t}$ is the distance from sensor j to the source target t and $\beta \geq 2$ is the signal decay exponent, which is assumed to be known (or estimated via training sequences [7]). This model can be found in many practical scenarios, such as in [8, 9].

We assume that the receiver has a Gaussian noise with p.s.d. $N_j, j \in \{1, \ldots, n\}$ within the baseband $[-\frac{\mathcal{W}}{2}, \frac{\mathcal{W}}{2}]$, where $\mathcal{W}$ is the bandwidth available on each link among sensors. We also assume that the radio frequency channels between neighboring sensors during period $\mathcal{T}$, are static, where $\mathcal{T}$ is the period required to perform the EF operation by each sensor. We consider that each sensor can adjust its communication cost (transmission power) such that a desired SNR S_0 (as defined for instance in [10]) is provided at the receiver to ensure a reliable communication. In this sense, the communication cost from an active sensor j to another active neighbor k is $f_c(d_{j,k}) = S_0 N_j d_{j,k}^{\alpha}(2^{B_j} - 1)/\mu$, where $d_{j,k}^{\alpha}$ is a channel gain and α ($2 < \alpha < 6$, depending on the WSN scenario) is the path-loss exponent [5, 11], B_j is the total number of bits per sample used to quantize samples at sensor j, and μ ($\mu > 0$) is a parameter that depends on the particular modulation scheme.

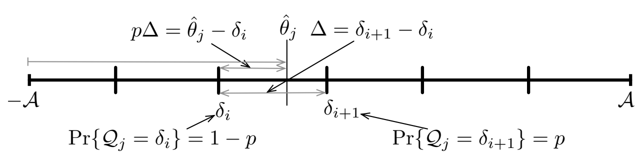

Fig. 2: The probabilistic quantization (PQ) method [12, 13].

Because of the strict bandwidth and energy limitations [14], each sensor is prevented from transmitting analog data, hence a local quantization is performed before transmission. Suppose that we wish to obtain a quantized message $\mathcal{Q}_j$ with B_j bits per sample for a local estimate $\hat{\theta}_j$ at sensor j, where we use a uniform quantization with $l = 2^{B_j}$ uniformly spaced quantization levels. Then, the threshold is given by a set $\delta = \{\delta_1, \ldots, \delta_l\}$ that follows that $\Delta = \delta_{i+1} - \delta_i = 2\mathcal{A}/(2^{B_j} - 1)$. Assuming that $\hat{\theta}_j$ is bounded within $[-\mathcal{A}, \mathcal{A}]$ as shown in Fig 2, then the variance of the quantization error when quantizing $\hat{\theta}_j$ is given by $\sigma^2_{q,j} = \Delta^2/4 = \mathcal{A}^2/(2^{B_j} - 1)^2 \leq 4\mathcal{A}^2/2^{2B_j}$, when a uniform and Probabilistic Quantizer (PQ), defined as $Q(\hat{\theta}_j) = \mathcal{Q}_j(\hat{\theta}_j, B_j)$ [12, 13] is used. In this quantization method, $\hat{\theta}_j$ is quantized as follows. If $\hat{\theta}_j$ lies between two quantization thresholds: δ_i and δ_{i+1} (see Fig 2), then $\hat{\theta}_j$ is quantized to the threshold δ_i with probability $1 - p$ and it is quantized to the threshold δ_{i+1} with probability p. Here, $p = (\hat{\theta}_j - \delta_i)/\Delta$ depends on the input distribution and it is chosen such that the quantization $\mathcal{Q}_j$ is unbiased. Thus it provides two important properties: the error of this quantization method has zero mean, $\mathbb{E}\{\mathcal{Q}_j\} = \hat{\theta}_j$; and the variance $\sigma^2_{q,j}$ can be bounded as, $\sigma^2_{q,j} \leq \mathcal{A}^2/2^{2B_j}$, where this inequality holds when $B_j \geq 1$. In this work, we are not considering more complex quantization methods such as vector quantization since a close-form expression for the variance of a realistic vector quantization is difficult to obtain.

2.1. Progressive Parameter Estimation

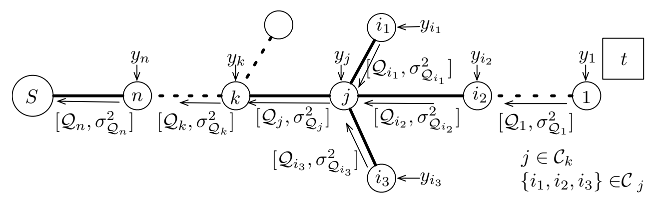

Fig. 3: Illustration of a multihop progressive estimation scheme.

Let us assume that the input to the j-th sensor (as shown in a simple example Fig. 3) consists of the sensor measurement y_j and the information $[\mathcal{Q}_i, \sigma^2_{\mathcal{Q}_i}], i \in \mathcal{C}_j$ that are received from all its children sensors $\{i_1, i_2, i_3\} \in \mathcal{C}_j$, where $\mathcal{C}_j$ is a set of children nodes of sensor j. We represent the quantized estimation that is received from sensor i by $\mathcal{Q}_i$ and its associated variance by $\sigma^2_{\mathcal{Q}_i}$. Then, the BLUE [5, 6] of θ in terms of y_j, $\mathcal{Q}_i$, and $\sigma^2_{\mathcal{Q}_i}$ at sensor j is given by:

$$\hat{\theta}_j = \left(\frac{b_j h_j^2}{\sigma^2_{y_j}} + \sum_{i \in \mathcal{C}_j} \frac{1}{\sigma^2_{\mathcal{Q}_i}} \right)^{-1} \left(\frac{b_j h_j y_j}{\sigma^2_{y_j}} + \sum_{i \in \mathcal{C}_j} \frac{\mathcal{Q}_i}{\sigma^2_{\mathcal{Q}_i}} \right) \tag{2}$$

where we assume a binary variable $b_j, j = 1, \ldots, n$ to assign the status of a sensor such that $b_j = 1$ refers to an active sensor and $b_j = 0$ refers to an inactive sensor. Then, the variance of $\hat{\theta}_j$, which is the MSE of the estimator, is given by:

$$\sigma^2_{\hat{\theta}_j} = \left(\frac{b_j h_j^2}{\sigma^2_{y_j}} + \sum_{i \in \mathcal{C}_j} \frac{1}{\sigma^2_{\mathcal{Q}_i}} \right)^{-1} \tag{3}$$

and the variance of the quantization error at sensor j is $\sigma^2_{q,j}$, then the variance of quantized estimation $\mathcal{Q}_j$ at sensor j is:

$$\sigma^2_{\mathcal{Q}_j} = b_j\sigma^2_{q,j} + \left(\frac{b_j h_j^2}{\sigma^2_{y_j}} + \sum_{i \in \mathcal{C}_j} \frac{1}{\sigma^2_{\mathcal{Q}_i}} \right)^{-1} \tag{4}$$

Consider first a simple case when only one child i of sensor j is available, that is, $|\mathcal{C}_j| = 1$, where $|\mathcal{C}_j|$ denotes the cardinality of $\mathcal{C}_j$, then (4) is given by:

$$\sigma^2_{\mathcal{Q}_j} = b_j\sigma^2_{q,j} + \left(\frac{b_j h_j^2}{\sigma^2_{y_j}} + \frac{1}{\sigma^2_{\mathcal{Q}_i}}\right)^{-1} = b_j\sigma^2_{q,j} + \frac{\sigma^2_{y_j}\sigma^2_{\mathcal{Q}_i}}{\sigma^2_{y_j} + \sigma^2_{\mathcal{Q}_i} b_j h_j^2} \tag{5}$$

Since (4) is a nonlinear recursion of $\sigma^2_{\mathcal{Q}_i}$, even if we apply the same argument recursively in a 1-D network (that is, $|\mathcal{C}_j| = 1$), it is hard to find a closed-form expression for $\sigma^2_{\mathcal{Q}_n}$, which is the final overall MSE based on all n sensors. It can be easily seen that $(\frac{\sigma_{y_j}}{\sqrt{b_j}h_j} - \sigma_{\mathcal{Q}_i})^2 \geq 0$, which gives $\frac{2\sigma_{y_j}\sigma_{\mathcal{Q}_i}}{\sqrt{b_j}h_j} \leq (\frac{\sigma^2_{y_j}}{b_j h_j^2} + \sigma^2_{\mathcal{Q}_i})$. Then, without loss of generality, (5) can be bounded [5, 10] as:

$$\sigma^2_{\mathcal{Q}_j} \leq b_j\sigma^2_{q,j} + \frac{\sigma^2_{y_j}}{4b_j h_j^2} + \frac{\sigma^2_{\mathcal{Q}_i}}{4} \tag{6}$$

And the equivalent form of (6) for $|\mathcal{C}_j| > 1$ of sensor j is:

$$\sigma^2_{\mathcal{Q}_j} \leq b_j\sigma^2_{q,j} + \frac{\sigma^2_{y_j}}{(1+|\mathcal{C}_j|)^2 b_j h_j^2} + \frac{1}{(1+|\mathcal{C}_j|)^2}\sum_{i\in\mathcal{C}_j}\sigma^2_{\mathcal{Q}_i} \tag{7}$$

The following recursion leads to a generalized form of the total estimation error due to all n sensors. To do this, we write an equivalent form of (7) for sensor k, where sensor k is the parent of sensor j ($k \leftarrow j \leftarrow i$), as shown in Fig. 3. In other words, $j \in \mathcal{C}_k$, that is, sensor j is now one of the children of sensor k, then (7) for sensor k becomes:

$$\sigma^2_{\mathcal{Q}_k} \leq b_k\sigma^2_{q,k} + \frac{\sigma^2_{y_k}}{(1+|\mathcal{C}_k|)^2 b_k h_k^2} + \frac{1}{(1+|\mathcal{C}_k|)^2}\sum_{j\in\mathcal{C}_k}\sigma^2_{\mathcal{Q}_j} \tag{8}$$

Using (7) and (8), we obtain:

$$\begin{aligned}\sigma^2_{\mathcal{Q}_k} \leq &\left(b_k\sigma^2_{q,k} + \frac{1}{(1+|\mathcal{C}_k|)^2}\sum_{j\in\mathcal{C}_k} b_j\sigma^2_{q,j}\right) + \\ &\left(\frac{\sigma^2_{y_k}}{(1+|\mathcal{C}_k|)^2 b_k h_k^2} + \frac{1}{(1+|\mathcal{C}_k|)^2}\sum_{j\in\mathcal{C}_k}\frac{\sigma^2_{y_j}}{(1+|\mathcal{C}_j|)^2 b_j h_j^2}\right) \\ &+ \frac{1}{(1+|\mathcal{C}_k|)^2}\sum_{j\in\mathcal{C}_k}\frac{1}{(1+|\mathcal{C}_j|)^2}\sum_{i\in\mathcal{C}_j}\sigma^2_{\mathcal{Q}_i}\end{aligned} \tag{9}$$

If sensor i is the leaf node, that is, $|\mathcal{C}_i| = 0$, then (4) becomes $\sigma^2_{\mathcal{Q}_i} = b_i\sigma^2_{q,i} + \frac{\sigma^2_{y_i}}{b_i h_i^2}$. In this case, equation (9) can be generalized as a sum of two terms:

$$\sigma^2_{\mathcal{Q}_n} \leq \sum_{j=1}^{n}\frac{b_j\sigma^2_{q,j}}{\mathcal{P}_j} + \sum_{j=1}^{n}\frac{\sigma^2_{y_j}}{\mathcal{P}_j(1+|\mathcal{C}_j|)^2 b_j h_j^2} \tag{10}$$

where $\sigma^2_{\mathcal{Q}_n}$ is the total MSE obtained at the sink node contributed by all n sensors and $\mathcal{P}_j = \prod_{l\in\Omega_j}(1+|\mathcal{C}_l|)^2$, where Ω_j is a set of all sensors in a single path between sensor j and the sink node S.

2.2. Optimization Problem

In order to transmit B_j bits per sample reliably from sensor j to its parent k, the minimum required communication cost must satisfy $f_c(d_{j,k}) = S_0 N_j d^{\alpha}_{j,k}(2^{B_j} - 1)/\mu < S_0 N_j d^{\alpha}_{j,k} 2^{B_j}$, based on the Shannon theory and a uniform quantization [5, 12]. Then, to optimize jointly three metrics: total distortion in estimation ($\sigma^2_{\mathcal{Q}_n}$), total communication cost, and number of bits per sample for each sensor, we formulate our problem as an optimization problem by assuming that the desired SNR S_0 is fixed for all sensors, that is:

$$\begin{aligned}\underset{\{b_j,b_k,B_j\}}{\text{minimize}} \quad & \sum_{j=1}^{n}\left(\frac{4b_j\mathcal{A}^2}{\mathcal{P}_j 2^{2B_j}} + \frac{\sigma^2_{y_j}}{\mathcal{P}_j(1+|\mathcal{C}_j|)^2 b_j h_j^2}\right) \\ \text{subject to} \quad & \sum_{j=1, j\in\mathcal{C}_k}^{n} b_j N_j^2 d^{2\alpha}_{j,k,} 2^{2B_j} \leq P_{\max} \\ & B_j \geq 1 \\ & b_j \leq b_k, j\in\mathcal{C}_k, A_{j,k} = 1 \\ & b_j \in \{0,1\}\end{aligned} \tag{11}$$

where $||.||_2$ ($norm$-2) is chosen in the sensor power vector so that we can penalize effectively higher communication cost links. The constraint $b_j \leq b_k$ ensures that a sensor j ($j \in \mathcal{C}_k$) is selected only when its parent k is selected, that is, it ensures a subtree $T \subset G$ from the selected sensors, rooted at the sink node. Inequality $B_j \geq 1$ ensures that at-least one bit per sample is allocated to each sensor measurement.

3. SENSOR SELECTION AND ROUTING ALGORITHM

3.1. Fixed-Tree Relaxation-Based Adaptive Quantization

This section deals with an algorithm that decouples f_c, which controls the communication cost and the estimation process, $\sigma^2_{\mathcal{Q}_n}$. For this, we first generate a Shortest Path Tree based on Communication Cost (SPT-CC) rooted at the sink node with $B_j = 1$ bit per sample for each measurement. Then, we store an edge-set $\{(j,k)\}$, where k is the parent of j, which is defined as a directed edge $j \to k$. Thus, the information for routing, $j \in \mathcal{C}_k$ in (11), is given. Finally, we rewrite the optimization problem to solve jointly the sensor selection (with variable b_j) and the bit allocation (with variable B_j) so that the routing structure used for selected sensors will be a subtree of the SPT-CC that has to be rooted at the sink node. We call this approach as Fixed-Tree Relaxation-based Adaptive Quantization (FTR-AQ) algorithm. Applying the EF scheme, a relaxed version of problem (11) is given by:

$$\begin{aligned}\underset{\{b^r_j,B_j\}}{\text{minimize}} \quad & \sum_{j=1}^{n}\left(\frac{4b^r_j\mathcal{A}^2}{\mathcal{P}_j 2^{2B_j}} + \frac{\sigma^2_{y_j}}{\mathcal{P}_j(1+|\mathcal{C}_j|)^2 b^r_j h_j^2}\right) \\ \text{subject to} \quad & \sum_{j=1, j\in\mathcal{C}_k}^{n} b^r_j N_j^2 d^{2\alpha}_{j,k} 2^{2B_j} \leq P_{\max} \\ & B_j \geq 1 \\ & b^r_j \leq b^r_k, j\in\mathcal{C}_k \\ & b^r_j \in [0,1]\end{aligned} \tag{12}$$

where b^r_j is the relax version of the variable b_j. Since problem (12) is a combinatorial optimization problem, in which the variable $b_j \in \{0,1\}$, and we can show as in [2] that this problem is generally an NP-hard. Notice that relaxed problem (12) is not a convex optimization problem because first inequality is a product of two convex functions, but it can be transformed into another problem by a *change of variables* and a *transformation* of the objective and constraints function. Let us define $r_j = 2^{B_j}/\sqrt{b_j^r}$, then (12) becomes:

$$\begin{aligned} \underset{\{b_j^r, r_j\}}{\text{minimize}} \quad & \sum_{j=1}^{n} \left(\frac{4\mathcal{A}^2 r_j^{-2}}{\mathcal{P}_j} + \frac{\sigma_{y_j}^2}{\mathcal{P}_j (1+|\mathcal{C}_j|)^2 b_j^r h_j^2} \right) \\ \text{subject to} \quad & \sum_{j=1, j\in\mathcal{C}_k}^{n} b_j^{r\,2} N_j^2 d_{j,k}^{2\alpha} r_j^2 \leq P_{\max} \\ & b_j^r \leq b_k^r,\ b_j^r \in [0,1],\ j \in \mathcal{C}_k \\ & r_j \geq 2, j = 1, \ldots, n \end{aligned} \tag{13}$$

This is now a Geometric Programing (GP) problem [15] since the objective is a posynomial function of b_j^r and r_j. The constraints are also expressed as posynomial inequalities, therefore it can be solve efficiently by GP programming (using interior-point method [15]).

We use the solution $\{b_j^{r*}, r_j^*\}$ of problem (13) to perform a suboptimal subset selection by sorting the optimal values $\{b_j^{r*}\}_{j=1}^n$ in descending order and selecting a subset $\mathcal{S}$ of K largest b_j^{r*}'s satisfying the power constraint inequality. B_j^*'s are obtained by $B_j^* = \frac{1}{2}\log_2(\frac{1}{2} r_j^{*2} b_j^{r*})$ that need to be rounded up to the nearest integer. Because of the subset selection, number of children $|\mathcal{C}_l|$ for each sensor in $\mathcal{S}$ need to be updated, therefore $\mathcal{P}_j = \prod_{l\in\Omega_j}(1+|\mathcal{C}_l|)^2$ becomes $\mathcal{P}_j^* = \prod_{l\in\Omega_j}(1+|\mathcal{C}_l^*|)^2$ with updated value $|\mathcal{C}_l^*|$. Then, denoting $\{\widehat{b}_j^r\}_{j=1}^n$ a set of binary values such that $\widehat{b}_j^r = 1$ if $j \in \mathcal{S}$ and $\widehat{b}_j^r = 0$ if $j \notin \mathcal{S}$, we have that:

$$L_{\text{FTR-AQ}} = \sum_{j=1}^{n} \left(\frac{4\widehat{b}_j^r \mathcal{A}^2}{\mathcal{P}_j^* 2^{2B_j^*}} + \frac{\sigma_{y_j}^2}{\mathcal{P}_j^* (1+|\mathcal{C}_j^*|)^2 \widehat{b}_j^r h_j^2} \right) \tag{14}$$

Further, it is noted that because of the relaxation and the constraint $b_j^r \leq b_k^r$ in (13), sorting the variables $\{b_j^{r*}\}$ makes a subtree T of SPT-CC, which is routed at the sink node.

3.2. Local Distributed Optimization Adaptive Quantization

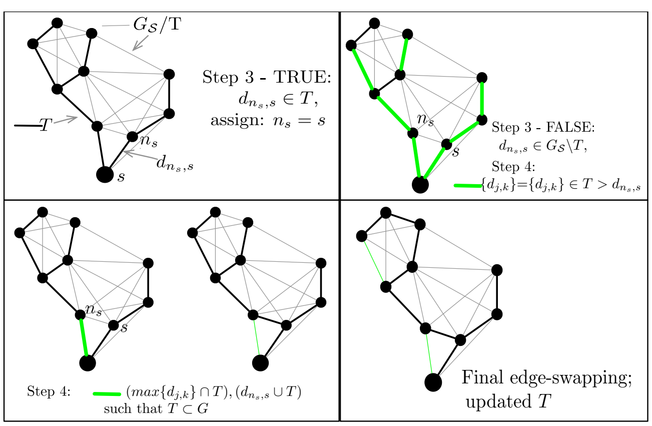

Fig. 4: Graphical representation of the edge-swap method [2]; Step 3 and Step 4 of **Algorithm 2**.

Finding a possible subset $\mathcal{S}$ of K selected sensors, their assigned bits per sample $B_j^*, j \in \mathcal{S}$, and an associated routing structure T from the solution of problem (13), can be improved by two operations: 1) performing an edge-swap method [2] on the subset $\mathcal{S}$ and 2) optimizing B_j^* locally and independently at each sensor. We call this process as a Local Distributed Optimization Adaptive Quantization (LDO-AQ) algorithm. First, we perform the edge-swap and then optimize B_j^* only for swapped edges. In order to perform this, we first define a subgraph $G_{\mathcal{S}}$, which is a graph obtained from restricting the graph G to the subset $\mathcal{S}$. Then, we perform swaps among the edges in T and the edges in $G_{\mathcal{S}}\backslash T$ in such a way that the new resulting tree after an update (swap) in the edges, remains a non-spanning tree of a graph G and that must be routed at the sink node. Steps of the procedure are given in **Algorithm 2** and then are illustrated in Fig. 4.

Algorithm 1 LDO-AQ Algorithm

Require: $\mathcal{S}, T, B_j^*, B_{\max}, K$

Initialization:

- $s = n+1$; $m = 1$; n_s= nearest sensor to s;

Step 1: find n_s
Step 2: if $m = K$, Stop; otherwise continue
Step 3: if $d_{n_s,s} \in T$ go to Step 6; otherwise continue
Step 4: find $\{d_{j,k}\} = \{f_c(d_{j,k})\} \in T > f_c(d_{n_s,s})$
remove largest edge, $T := T\backslash \max\{d_{j,k}\} \ni T \subset G$;
update, $|\mathcal{C}_k^*| := |\mathcal{C}_k^*| - 1$;
and add edge $d_{n_s,s}$, $T := T \cup d_{n_s,s} \ni T \subset G$;
update, $|\mathcal{C}_s^*| := |\mathcal{C}_s^*| + 1$;
power gain, $P_{\text{gain}} = f_c(\max\{d_{j,k}\}) - f_c(d_{n_s,s})$;
Step 5: update $B_{n_s}^*$ for the edge $d_{n_s,s}$ such that $P_{\text{gain}} \to 0$
for $B_{n_s}^*$=1 to $B_{\max}$ **do**
  $P_{\text{gain}} = N_j \max\{d_{j,k}\}^\alpha 2^{B_j^*} - N_{n_s} d_{n_s,s}^\alpha 2^{B_{n_s}^*}$;
  if $P_{\text{gain}} < 0$ **then**
    stop and update: $B_{n_s}^* = B_{n_s}^* - 1$;
  end if
end for
assign: $n_s = s$; $m = m+1$ go to Step 1
Step 6: assign $n_s = s$; $m = m+1$ go to Step 1

The idea given by these steps is the following: first, we assign $s = n+1$ (identity of the sink node) and an index $m = 1$, then find the nearest sensor n_s to s based on the communication cost and the solution B_j^* of problem (13). Find, if the edge $d_{n_s,s} \in T$ is true (in **Step 3**), then there is no swap, go to **Step 6** to assign $n_s = s$, increase m by one, and then return to **Step 1** to repeat the process. On the other hand, if $d_{n_s,s} \notin T$, then find the list $\{d_{j,k}\}$ of all the edges of T that are with the larger communication cost than the communication cost of the edge $d_{n_s,s}$. Find the largest $d_{j,k}$ from the list $\{d_{j,k}\}$ such that while swapping it with $d_{n_s,s}$ ensures that T is still routing at the sink node. Update T and number of children $|\mathcal{C}^*|$ by removing the edge $d_{j,k}$ and adding the edge $d_{n_s,s}$ and calculate a power gain $P_{\text{gain}} = N_j \max\{d_{j,k}\}^\alpha 2^{B_j^*} - N_{n_s} d_{n_s,s}^\alpha 2^{B_{n_s}^*}$. In some cases, P_{gain} is very large, which can be minimized by optimizing $B_{n_s}^*$ locally (since, now measurement flows from n_s to s and $d_{n_s,s} < d_{j,k}$) as in the **Step 5**. Increasing $B_{n_s}^*$ from its current value tends to reduce P_{gain} toward zero and then a higher $B_{n_s}^*$ provides less quantization error. After optimizing $B_{n_s}^*$, assign $n_s = s$, increase the index m by one, and return to **Step 1** to repeat the process until we scan all the edges in T, that is, $m = K$. Notice that, each edge in T is scanned only once. Finally, we calculate, in this case (14), let us call this solution $L_{\text{LDO-AQ}}$.

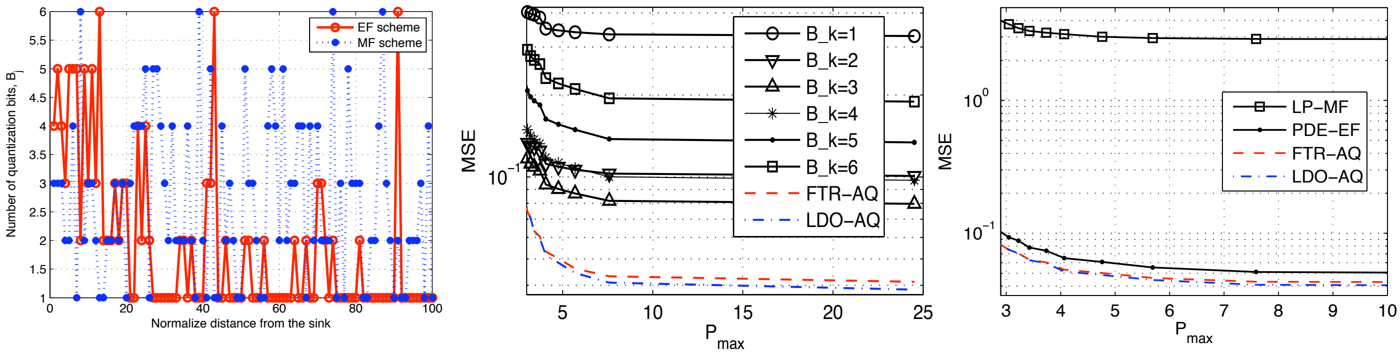

Fig. 5: (a) Allocated B_j vs. the normalized Euclidean distance from a sensor to the sink node for $P_{\max} = 10$, (b) MSE performance, when activating a subset of sensors for a given range of power budget $P_{\max}$ and comparing their performance with the fixed quantization algorithm; given quantization levels from $B_j = 1$ to $B_j = 6$, (c) MSE performance of the proposed and related algorithms for a given range of $P_{\max}$. Related algorithms are: LP-MF [4]; and PDE-EF [5].

4. SIMULATION RESULTS

We consider a WSN with $n = 100$ randomly deployed sensors. We assume $\alpha = 4$ and $N_j = 1 \forall j$ in our communication cost model $f_c(d_{j,k}) < S_0 N_j d_{j,k}^{\alpha} 2^{B_j}$. The measurement gain in our example is assumed $h_j = 1/d_{j,t}^{\beta}$ with $\beta = 2$. Without loss of generality, we assume that $\sigma_{z_j}^2 = \sigma_{y_j}^2 = 1 \forall j$ and $\mathcal{A} = 1$. We tested our algorithms using 100 different network topologies and a range of power constraint values to each network topology are provided. All simulation results are averaged over these 100 different network topologies. In Fig. 5(a), we show an example of bits per sample distribution using the MF and EF schemes. It can be noticed that in the MF scheme the bit distribution (marked by blue stars) is nearly uniform (that is, all levels provided), whereas in the EF scheme sensors close to the sink node (generally small number of hops, that is, small normalizing distance) allows more bits per sample in order to maintain small total estimation error.

A comparison between fixed and adaptive quantization is shown in Fig. 5(b), where six different cases of fixed quantization are taken from $B_j = 1$ to 6 in the fixed tree relaxation-based algorithm [2]. Then, these six simulation results are compared with the proposed approaches. It can be seen that our adaptive quantization algorithms perform superior to the fixed quantization. We consider two heuristic algorithms to compare with our proposed algorithms, which are LP-MF [4], and PDE-EF [5]. Algorithm LP-MF considers routing, but uses the MF scheme, PDE-EF uses the EF scheme, but does not optimize jointly all three metrics. An MSE from our FTR-AQ is given by (14) and an improved MSE for our LDO-AQ is obtained by **Algorithm 2**. Fig. 5(c) shows MSE performances for LP-MF, PDE-EF, FTR-AQ, and LDO-AQ algorithms for a given power budget range $P_{\max} = 3$ to 10. We can observe that our approaches outperform the other related algorithms cited to compare in this work. The performance of PDE-EF is close to our approaches since it uses the EF scheme.

5. CONCLUSION

Most of the recent solutions that have been proposed, try to reduce the problem to a subset selection, ignoring the optimization of the routing structure as well as source coding effectively. However, optimizing the routing structure is an important metric in the problem since in general, transmitting an information that is faraway from the sink node requires more energy than one that is close to it. Source coding also plays an important role as can be seen in Fig. 5(a) that a sensor faraway from the sink node requires few bits/sample in quantizing it's measurement when using the EF scheme.

6. REFERENCES

[1] Santosh Shah and B. Beferull-Lozano, "In-network iterative distributed estimation for power-constrained wireless sensor networks," DCOSS, pp. 239-246, 2012 *IEEE 8th Inter. Conf. on Distributed Comp. in Sensor Systems, 2012*.

[2] Santosh Shah and Baltasar Beferull-Lozano, "Power-aware joint sensor selection and routing for distributed estimation: a convex optimization approach," DCOSS, pp. 230-238, 2012 *IEEE 8th Inter. Conf. on Distributed Comp. in Sensor Systems, 2012*.

[3] Bang Ye Wu and Kun-Mao Chao, *Spanning Trees and Optimization Problems*, Chapman & Hall/CRC Press, USA, 2004.

[4] J. Li and G. AlRegib, "Maximizing network lifetime for estimation in multihop wireless sensor networks," *Computer Comm. and Networks*, IC-CCN 08, pp. 1-6, August 2008.

[5] Y. Huang, Y. Hua, "Multi-hop progressive decentralized estimation in wireless sensor networks," *IEEE Sig. Process. Lett.*, vol. 14, p. 1004-07 , 2007.

[6] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, volume I, Pentice-Hall, Inc, 2007.

[7] Dan Li, Wong K.D., Yu Hen Hu, and Sayeed, A.M., "Detection, classification, and tracking of targets," *IEEE Signal Processing Magazine*, vol. 19, no. 2, pp. 17-29, Mar 2002.

[8] Feng Zhao and Leonidas Guibas, *Wireless Sensor Networks: An Information Processing Approach*, (The Morgan Kaufmann Series in Networking). Morgan Kaufmann, July 2004.

[9] Kim, Y.H. and Ortega, A., "Quantizer design and distributed encoding algorithm for source localization in sensor networks," *Information Processing in Sensor Networks*, pp. 231-238, 15 April 2005.

[10] Y. Huang and Y. Hua, "Energy planning for progressive estimation in multihop sensor networks," *IEEE Tran. on Signal Proc.*, vol. 57, no. 10, pp. 4052-4065, October 2009.

[11] Zhi Quan, Kaiser, W. J., and Sayed, A. H., "Innovations diffusion: a spatial sampling scheme for distributed estimation and detection," *IEEE Tran. on Sig. Proc.*, vol. 57, no. 2, pp. 738-751, Feb. 2009.

[12] J. J. Xiao, S. Cui, Z-Q. Luo, and A. J. Goldsmith, "Power scheduling of universal decentralized estimation in sensor networks," *IEEE Tran. on Signal Proc.*, vol. 54, no. 2, pp. 413-422, Feb. 2006.

[13] Xiao, J. J. and Z. Q. Luo, "Decentralized estimation in an inhomogeneous sensing environment," *IEEE Transactions on Information Theory*, vol. 51, no. 10, pp. 3564-3575, Oct. 2005.

[14] A. Ribeiro and G. Giannakis, "Bandwidth-constrained distributed estimation for wireless sensor networks - part I Gaussian case," *IEEE Tran. on Signal Processing*, vol. 54, no. 3, pp. 1131-1143, March 2006.

[15] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, March 2004.