Power Constrained Linear Estimation in Wireless Sensor Networks with Correlated Data and Digital Modulation

Muhammad Hafeez Chaudhary*, *Student Member, IEEE,* and Luc Vandendorpe, *Fellow, IEEE*

***Abstract*—We study the problem of joint quantization and power allocation in wireless sensor networks where spatially distributed sensors observe a Gaussian random source, quantize the resulting noisy observations, and transmit over orthogonal fading channels to a remote fusion center (FC). The role of the FC is to reconstruct the source with minimal distortion using linear minimum mean square error estimation rule. In this paper, we undertake the design of joint quantization and power allocation based on the following optimization problem: minimize the reconstruction distortion for a given total network power consumption. To address this problem, at each sensor node uniform scalar quantization is assumed. Moreover, assuming pseudo-quantization noise model we show that the problem can be solved using a block-coordinate descent type algorithm which iteratively optimizes the quantization bits and the power allocations. The algorithm takes into account the spatial correlation, the observation noise, and the channel quality of the sensors. Numerical and simulation examples corroborate the analytical results. The examples illustrate that the proposed design holds a considerable performance gain compared to a quantization scheme based on the uniform power allocation.**

***Index Terms*—Digital modulation, orthogonal multiple access channel, parameter estimation, quantization, resource management, spatial correlation, wireless sensor networks.**

I. Introduction

Wireless sensor networks (WSNs) consist of spatially distributed sensors that cooperatively monitor physical or environmental conditions. The sensor networks are characterized by the limited availability of energy, bandwidth, and computational power. Our objective is to reconstruct the underlying signal subject to resource constraints so that the overall distortion for instance, mean square error (MSE) be minimized. We consider a system in star topology where sensors transmit quantized version of their noisy observations via some orthogonal multiple access scheme (e.g., TDMA or FDMA) to a central processing unit called fusion center (FC) which produces a global picture of the physical phenomenon. The sensors have partial and correlated observations of the source. The correlation exists where sensors measure data in the same geographical location. In addition, observation noise and communication channel may not have the same conditions across all sensors. Thus independent quantization and transmission of the observations is not an optimal strategy.

For estimation of a Gaussian source under mean-squared distortion measure, in the information theoretic perspective, it has been shown in [1] that the digital transmission is optimal in a WSN observing the source where sensors have orthogonal channels to the FC. This result combined with the advantages of the digital communication scheme—such as modularity and robustness, among others—motivate us to study the problem of adaptive joint quantization and power allocation (AJQPA) in WSNs.

In [2] a quantization and power allocation scheme is proposed where a parameter is estimated based on the best linear unbiased estimation (BLUE) rule which does not exploit the spatial correlation. Krasnopeev and his colleagues in [3] consider quantization of the observations where observation noise is correlated across sensors. In [4], the authors propose a tradeoff between the number of active sensors and the quantization bits of each sensor with the objective of minimizing the estimation MSE subject to a network-wide rate constraint. The work in [5] and [6] considers estimation of a source by the BLUE rule with binary modulated transmission of the quantized observations. Some other related works appear in [7]–[10]. More recent work in [11] presents the idea wherein sensors quantize their observation using identical one-bit quantizers based on the minimization of Cramér–Rao lower bound (CRLB) on the estimation error. Another related work appears in [12], which deals with the one-bit quantizer design for estimation in WSNs. For a vector parameter estimation in WSNs, [13] presents a joint scheme to compress the multidimensional observation at each sensor to a scalar value and then design a one-bit quantizer to quantize the compressed value with a target to minimize the CRLB distortion measure. These works do not consider the transmission of the observations over non-ideal channels. The exposition in [14] studies the problem of optimal network size in the frame work of information theoretic source coding. In the aforementioned works, an unknown deterministic parameter is estimated by a set of distributed sensor nodes based on the quantized sensor observations. However, therein they do not exploit the spatial correlation, and in some cases assume ideal communication channels or consider homogeneous sensor networks.

In [15]–[19] estimation schemes are proposed which exploit the spatial correlation in WSNs, but do not consider the

The authors are with the Institute of Information and Communication Technologies, Electronics and Applied Mathematics (ICTEAM), Université catholique de Louvain, 1348 Louvain-La-Neuve, Belgium (email: {Muhammad.Chaudhary, Luc.Vandendorpe}@uclouvain.be).

The authors would like to thank ARC SCOOP project, Walloon region framework program COSMOS/TSARINE, and EU project FP7 NEWCOM++ for the financial support and the scientific inspiration.

quantization of the sensor observations. In fact sensors simply amplify and forward their noisy analog observations to the FC. In this work, we present a joint design for quantization and power allocation in WSNs which takes into account the spatial correlation and cross-correlations of the observations, the quality of the observations, and the quality of the communication channel from sensors to the FC. For solution to the underlying optimization problem, we propose a block-coordinate descent type algorithm which iteratively optimizes the quantization bits and the transmit powers while minimizing the distortion function.

At each sampling instant, the sensors quantize the observations employing the uniform scalar quantization scheme, encode the quantization indices, and transmit the resulting bits to the FC. The FC reconstructs the underlying source based on a linear minimum mean square error (LMMSE) estimation rule. The quantization and power allocation scheme targets minimization of the reconstruction distortion subject to a constraint on the total transmit power of the sensors. To this end, an AJQPA scheme is presented in [20] where we do not make any simplifying assumption be it on the quantization noise or on the contribution of the channel errors to the total reconstruction distortion. That scheme is computationally expensive and becomes intractable with the network size. Despite the complexity issue, the AJQPA scheme discussed there under reasonably simple scenarios (e.g., small network size) can serve as a baseline approach for performance comparison and can be used as a tool to check the validity of the simplifying assumptions of other schemes. In our pursuit to simplify the AJQPA problem, we use in the current work the pseudo-quantization noise model [21], [22]. Based on that, the reconstruction distortion is formed in two steps: Firstly, assuming error-free transmission of the quantized sensor observations to the FC we characterize the distortion; subsequently the contribution of the channel errors to the distortion is quantified and added to form the total distortion. Based on this distortion function, to solve the AJQPA problem we propose a block-coordinate descent type algorithm which scales well with the network size. The proposed design leads to better performance compared to a quantization scheme based on uniform power allocation to the sensors.

Compared to the existing literature, the contributions of this work lie in the following aspects:

- The proposed design for AJQPA in the sensor networks jointly exploits the spatial correlation, the observation noise, and the channel quality of the sensors. In WSN applications where for example, the underlying source is acoustic pressure, heat, or chemical concentration, the sensor observations are better characterized by spatial correlation based observation models as in [15], [17]–[19]. However, the existing works on quantization for estimation in the sensor networks, like [2], assume that the underlying source is spatially invariant and deterministic, but otherwise unknown. Thus they do not consider the spatial variations of the source and do not incorporate any a priori knowledge about the source.
- The proposed AJQPA scheme gives solution which is markedly simpler than the scheme in [20] and gives distortion performance which compares favorably with that of [20]. Compared to the work by Xiao and his colleagues [2], where a joint quantization and power allocation scheme is proposed, the underlying system model (as mentioned before), the optimization problem, and the resulting solution are notably different for the AJQPA scheme presented in this paper. The joint quantization and power allocation problem in [2] under certain assumptions—the given target bit-error probability for the sensors, the choice of the sum of the squares of the sensor transmit powers as the objective function, and so on—is transformed into an optimization problem over the quantization bits only. The sensor transmit powers are obtained as a function of the quantization bits for a fixed bit-error probability. However, in the current work we jointly optimize the quantization bits and the transmit powers (and thus also optimize the bit-error probability).

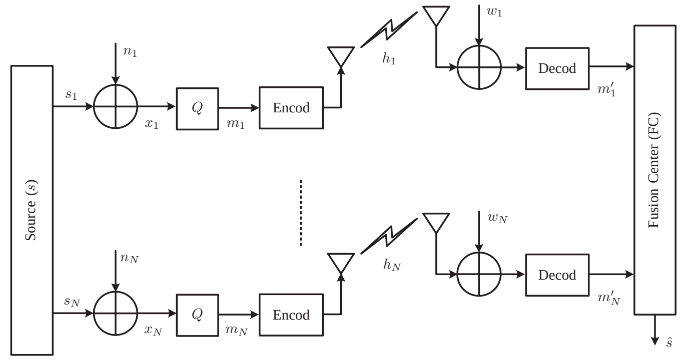

Fig. 1: Block diagram of the system.

The rest of the paper is organized as follows. Section II describes the network setup under consideration. In Sections III, we formulate the AJQPA problem and outline its solution. To substantiate the analytical findings, Section IV presents some numerical and simulation examples. Finally, Section V provides some concluding remarks.

II. System Model

Consider the system model shown in Fig. 1, where N spatially distributed sensors observe an unknown Gaussian random source $s \sim \mathcal{N}\big(0, \sigma_s^2\big)$. Each sensor, say sensor i, has a partial observation $s_i \sim \mathcal{N}\big(0, \sigma_{s_i}^2\big)$ of the source s. The observation is corrupted by additive noise $n_i \sim \mathcal{N}\big(0, \sigma_{n_i}^2\big)$ such that the noisy observation at the sensor is given by

$$x_i = s_i + n_i, \quad i = 1, \ldots, N, \tag{1}$$

where n_i is independent across sensors, and is also independent of s and s_i's. We assume that the sensors employ the uniform scalar quantization scheme to quantize their observations. We also assume that the sensors transmit the quantization indices to the FC via orthogonal channels; where the channels experience flat fading independent over time and across sensors. The fading channels, with gain factors $g_i = |h_i|$ for all i, are assumed to be perfectly known at the FC and do not change during the estimation of each observation sample.

The $w_i \sim \mathcal{CN}\big(0, 2\sigma_{w_i}^2\big)$ denotes the receiver noise, which is independent across the sensors, and is also independent of s, s_i's, and n_i's. In Fig. 1, the *Encod* block at each sensor node performs two functions: It encodes the indices of the quantized values according to some labeling rule for example, natural binary code, and then modulates the resulting bits (also called here quantization bits) using some digital modulation scheme for example, BPSK modulation. On the other hand, each *Decod* block at the FC performs the converse functions of the corresponding *Encod* block—that is, demodulation and mapping of the received bits to the reconstruction values.

We assume that the source s, the observation s_i at sensor i, and the observation s_j at sensor j are jointly Gaussian distributed having zero mean and covariances $\mathrm{Cov}\{s, s_i\} = \sigma_s \sigma_{s_i} \rho_{s,i}$, $\mathrm{Cov}\{s, s_j\} = \sigma_s \sigma_{s_j} \rho_{s,j}$, and $\mathrm{Cov}\{s_i, s_j\} = \sigma_{s_i} \sigma_{s_j} \rho_{i,j}$ for all i and j. Note that $\rho_{s,i}$ specifies the correlation between s and s_i, and $\rho_{i,j}$ specifies the correlation between s_i and s_j. Moreover, we assume that the samples of each of s, s_i, n_i, and w_i are independent in time. The spatial correlation $\rho_{s,i}$ and cross-correlation $\rho_{i,j}$ coefficients can be modeled as follows:

$$\rho_{s,i} = e^{-(d_{s,i}/\theta_1)^{\theta_2}}, \quad \rho_{i,j} = e^{-(d_{i,j}/\theta_1)^{\theta_2}}, \tag{2}$$

respectively. Where $d_{s,i}$ is distance between the source s and the sensor i, $d_{i,j}$ is distance between the sensors i and j, $\theta_1 > 0$ is a range parameter which controls how fast the correlation decays with the distance, and $0 < \theta_2 \leq 2$ is called a smoothness parameter [15].

When the sensor observations x_i's are available at the FC then the optimal estimator in the minimum MSE sense is the conditional mean of s given x_i's that is, $\hat{s}_0 = \mathbb{E}[s|x_i, \forall i]$, where $\mathbb{E}$ denotes the mathematical expectation. Under the jointly Gaussian assumption of s and x_i's, the conditional mean estimator turns out to be linear and is called the LMMSE estimator which can be written as $\hat{s}_0 = \mathbf{c}^T(\mathbf{C_s}+\mathbf{C_n})^{-1}\mathbf{x}$ with the associated MSE distortion given by

$$D_0 = \sigma_s^2 - \mathbf{c}^T(\mathbf{C_s} + \mathbf{C_n})^{-1}\mathbf{c}, \tag{3}$$

where $\mathbf{c} = \mathbb{E}[\mathbf{x}s]$, $\mathbf{C_s} = \mathbb{E}[\mathbf{s}\mathbf{s}^T]$, and $\mathbf{C_n} = \mathbb{E}[\mathbf{n}\mathbf{n}^T]$ with $\mathbf{x} = [x_1, \ldots, x_N]^T$, $\mathbf{s} = [s_1, \ldots, s_N]^T$, and $\mathbf{n} = [n_1, \ldots, n_N]^T$ [23]. Note that $[.]^T$ denotes the matrix–vector transpose operation.

Remark 1: The estimator which achieves the distortion D_0 is called a clairvoyant estimator and it can be used as a performance benchmark, as the distortion achieved by any estimator designed to minimize the MSE distortion measure is lower bounded by D_0.

We can view the quantization function Q_i at sensor i as a mapping, which maps the observation x_i to one of the finite rational numbers $\{m_{i,1}, \ldots, m_{i,M_i}\}$ as follows:

$$\begin{aligned} &Q_i : x_i \mapsto m_i \\ &m_i = m_{i,k}, \quad \text{for } u_{i,k} < x_i \leq u_{i,k+1},\, k = 1, \ldots, M_i, \end{aligned} \tag{4}$$

where $u_{i,k}$'s are quantization interval $\varDelta_i$ boundaries and $m_{i,k}$'s are quantization values (also called representation or reconstruction values). The sensor i encodes the index k corresponding to the value $m_{i,k}$ and transmits the resulting bits to the FC. The encoding of the quantization indices does not consider entropy coding that is, a fixed length coding scheme is used. Thus, we require $L_i = \log_2(M_i)$ bits to encode the M_i indices, which we call as the quantization bits. Moreover, in this work, we do not consider channel coding.

For the uniform scalar quantization, the quantization boundaries can be written in terms of the quantization interval $\varDelta_i$ as $u_{i,k} = (2k - 2 - M_i)\varDelta_i/2$ for $k = 2, \ldots, M_i$ with $u_{i,1}$ and u_{i,M_i+1} denoting the greatest lower bound and the lowest upper bound on x_i, respectively. The corresponding quantization values are given by $m_{i,k} = (2k - 1 - M_i)\varDelta_i/2$ for $k = 1, \ldots, M_i$. As x_i is a zero-mean Gaussian distributed random variable, therefore for some reasonably large value of W we have $p(|x_i| \geq W) \approx 0$. As a consequence we can assume that $u_{i,1} = -W$ and $u_{i,M_i+1} = W$. In this particular case, the quantization interval size $\varDelta_i$ can be given by $\varDelta_i = 2W/(2^{L_i} - 1) \approx 2W/2^{L_i}$.

III. Quantization and Power Allocation: Proposed Approach

Let D_t be the estimation distortion of the underlying source at the FC. The quantization and power allocation scheme is based on minimization of the distortion subject to a constraint on the total network power consumption as follows:

$$\begin{aligned} &\underset{L_i, P_i, \forall i}{\text{minimize}}\; D_\mathrm{t}(L_i, P_i; i = 1, \ldots, N) \\ &\text{subject to} \sum_{i=1}^{N} P_i \leq P_\mathrm{t},\, L_i \in \mathbb{Z}_+,\, P_i \in \mathbb{R}_+, \quad \forall i. \end{aligned} \tag{5}$$

Where $D_\mathrm{t}(L_i, P_i; i = 1, \ldots, N)$ is the distortion function; L_i and P_i are the quantization bits and the transmit power of sensor i, respectively; $\mathbb{Z}_+$ and $\mathbb{R}_+$ denote the set of positive integers and the set of positive real numbers, respectively. In (5), the constraint on the total power consumption enables a fair comparison between the networks of different sizes. Moreover, putting a cap on the total power consumption limits interference with the neighboring networks, conserves energy, and also makes sense from the view point of global energy efficiency [24].

In the sequel, we solve the optimization problem (5) based on an approximation of the distortion function. This approximation is obtained by assuming the pseudo-quantization noise model, which is a simplified model of the quantization process [21], [22]. We use thus obtained distortion function as a surrogate for D_t and solve the problem (5).

A. Preliminaries and the Problem Solution

The quantization noise is an error or a distortion introduced by the quantizer and is defined as a difference between the quantizer output and the input that is, $q_i := m_i - x_i$ for all i, which is bounded as $-\varDelta_i/2 \leq q_i \leq \varDelta_i/2$. In general, the quantization noise q_i is a non-zero mean random variable which is not uniformly distributed, and is correlated with the $\{q_j\}_{\forall j \neq i}$ and the inputs $\{x_i\}_{i=1}^N$. However, it has been shown in [21] that when quantizing the correlated Gaussian distributed variables $\{x_i\}_{i=1}^N$ with uniform quantizers, surprisingly the quantization noises $\{q_i\}_{i=1}^N$ are almost uncorrelated and independent from each other as well as from the inputs $\{x_i\}_{i=1}^N$ for extremely rough quantization (values of the quantization step size Δ_i up to σ_{x_i} or more) and for very high correlation coefficients (but strictly less than 1) between any x_i and x_j. Moreover, all of the quantization noises $q_1, \ldots, q_N$ can be approximated as zero-mean and uniformly distributed random variables between $\pm\Delta_1/2, \ldots, \pm\Delta_N/2$ with corresponding variances $\Delta_1^2/12, \ldots, \Delta_N^2/12$, respectively; that is, the quantization noise variance for sensor i is

$$\sigma_{q_i}^2 = \frac{W^2}{3(2^{L_i}-1)^2}, \quad \forall i. \tag{6}$$

Consequently, the quantization noises $\{q_i\}_{i=1}^N$ can be modeled as uniformly distributed and statistically independent of each other and of the inputs $\{x_i\}_{i=1}^N$. In what follows, we use this quantization noise model to design the AJQPA scheme. Therein, to form the estimate of the source s and to characterize the reconstruction distortion, we adopt a two-step procedure: In the first step, we assume that the quantized signals from sensors arrive error free at the FC; then in the second step, we incorporate the contribution of the bit errors caused by the channel. This kind of approach to estimation over noisy channels is used in the literature for instance, see [2], [6].

The FC employs an LMMSE estimator to reconstruct the source from the quantized signals of the sensors. Let m_i' be the quantized message received through the actual channel and m_i be the quantized message received error free at the FC. We can write estimates $\hat{s}$ and $\hat{s}'$ as follows:

$$\hat{s} = \sum_{i=1}^N \alpha_i m_i, \quad \hat{s}' = \sum_{i=1}^N \alpha_i m_i', \tag{7}$$

where α_i's are the LMMSE weighting coefficients. The total mean-squared reconstruction distortion at the FC that is, MSE of $\hat{s}'$ with respect to the actual parameter s can be defined as

$$\begin{aligned} D_{\mathrm{t}} &= \mathbb{E}_{\{s,s_i,n_i,w_i|h_i,\forall i\}}\left[(\hat{s}'-s)^2\right] \\ &= \underbrace{\mathbb{E}[(\hat{s}-s)^2]}_{:=D} + \underbrace{\mathbb{E}[(\hat{s}'-\hat{s})^2]}_{:=D_{\mathrm{c}}} + \underbrace{2\mathbb{E}[(\hat{s}'-\hat{s})(\hat{s}-s)]}_{:=D_{\mathrm{m}}}, \end{aligned} \tag{8}$$

where D accounts for the distortion arising from the spatial correlation, and the quantization and observation noises; D_{c} accounts for the distortion from the channel errors; and D_{m} characterizes the mutual term. Based on the Cauchy–Schwarz inequality we can upper bound the distortion D_{t} as follows:

$$\begin{aligned} D_{\mathrm{t}} &\le \mathbb{E}[(\hat{s}-s)^2] + \mathbb{E}[(\hat{s}'-\hat{s})^2] + 2\sqrt{\mathbb{E}[(\hat{s}'-\hat{s})^2]\mathbb{E}[(\hat{s}-s)^2]} \\ &= \left(\sqrt{D}+\sqrt{D_{\mathrm{c}}}\right)^2 \le 2\left(D+D_{\mathrm{c}}\right) = 2\tilde{D}_{\mathrm{t}} \end{aligned} \tag{9}$$

where $\tilde{D}_{\mathrm{t}} = D + D_{\mathrm{c}}$. To obtain (9) we have used $\left(\mathbb{E}[xy]\right)^2 \le \mathbb{E}\left[x^2\right]\mathbb{E}\left[y^2\right]$ in the first inequality and $\left(\sum_{k=1}^K D_k\right)^2 \le K\sum_{k=1}^K D_k^2$ in the second inequality. In the subsequent development for the AJQPA design, we target minimization of $\tilde{D}_{\mathrm{t}}$. To this purpose, first we characterize the distortion D and subsequently the distortion D_{c}.

The distortion D can be viewed as the MSE of the estimate $\hat{s}$ in comparison to the original signal s assuming that the quantized signals m_i's are perfectly received at the FC. In the following we determine the weighting factors α_i's such that the distortion D be minimize. The distortion $D = \mathbb{E}\left[(\hat{s}-s)^2\right]$ can be characterized as follows:

$$\begin{aligned} D = \sigma_s^2 + \sum_{i=1}^N \alpha_i^2\left(\sigma_i^2+\sigma_{q_i}^2\right) + \\ \sum_{i=1}^N\sum_{j\neq i}^N \alpha_i\alpha_j\sigma_{s_i}\sigma_{s_j}\rho_{i,j} - 2\sum_{i=1}^N \alpha_i\sigma_s\sigma_{s_i}\rho_{s,i}, \end{aligned} \tag{10}$$

where $\sigma_i^2 = \sigma_{s_i}^2 + \sigma_{n_i}^2$. For $i = 1, \ldots, N$, by taking derivative of (10) with respect to α_i and setting it equal to zero, we get the following expression for the weighting coefficients that minimizes the distortion D:

$$\alpha_i = \frac{\rho_{s,i}\sigma_s\sigma_{s_i} - \sum_{j\neq i}^N \beta_j\gamma_j\sigma_{s_i}\sigma_{s_j}\rho_{i,j}}{\sigma_i^2+\sigma_{q_i}^2} = \beta_i\gamma_i. \tag{11}$$

Where β_i and γ_i are, respectively, defined as

$$\gamma_i = \frac{1}{\sigma_i^2+\sigma_{q_i}^2}, \quad \forall i, \tag{12}$$

$$\beta_i = \rho_{s,i}\sigma_s\sigma_{s_i} - \sum_{j\neq i}^N \beta_j\gamma_j\sigma_{s_i}\sigma_{s_j}\rho_{i,j}, \quad \forall i. \tag{13}$$

From (6) and (12), we can express L_i as a function of γ_i as follows:

$$L_i = \log_2\left(1+\sqrt{\frac{W^2\gamma_i}{3\left(1-\gamma_i\sigma_i^2\right)}}\right). \tag{14}$$

Eq. (13) forms a set of N simultaneous linear equations which in the matrix–vector notation can be written as $\beta = \left(\mathbf{C}_\gamma\tilde{\mathbf{C}}_{\mathbf{s}}+\mathbf{I}\right)^{-1}\mathbf{c}$; where $\beta = [\beta_1, \ldots, \beta_N]^T$, $\mathbf{C}_\gamma = \mathrm{diag}\left(\gamma_1, \ldots, \gamma_N\right)$, and $\tilde{\mathbf{C}}_{\mathbf{s}} = \mathbf{C}_{\mathbf{s}} - \mathrm{diag}\left(\sigma_{s_1}^2, \ldots, \sigma_{s_N}^2\right)$. Note that $\mathbf{c} = \mathbb{E}[\mathbf{x}s]$ and $\mathbf{C}_{\mathbf{s}} = \mathbb{E}\left[\mathbf{s}\mathbf{s}^T\right]$ with $\mathbf{x} = [x_1, \ldots, x_N]^T$ and $\mathbf{s} = [s_1, \ldots, s_N]^T$. It can be shown that the vector $\alpha = [\alpha_1, \ldots, \alpha_N]^T$ can be written as

$$\alpha = \mathbf{C}_\gamma\beta = \left(\tilde{\mathbf{C}}_{\mathbf{s}}+\mathbf{C}_\gamma^{-1}\right)^{-1}\mathbf{c}. \tag{15}$$

With (11) the distortion in (10) simplifies to

$$\begin{aligned} D &= \sigma_s^2 - \sum_{i=1}^N \alpha_i\sigma_s\sigma_{s_i}\rho_{s,i} \\ &= \sigma_s^2 - \alpha^T\mathbf{c} = \sigma_s^2 - \mathbf{c}^T\left(\tilde{\mathbf{C}}_{\mathbf{s}}+\mathbf{C}_\gamma^{-1}\right)^{-1}\mathbf{c}, \end{aligned} \tag{16}$$

where note that the components of $\mathbf{C}_\gamma$ that is, γ_i is function of L_i for all i (c.f., (6) and (12)).

Note that the coefficients α_i's are optimal for the estimate $\hat{s}$ in the MMSE sense. We use the same weighting coefficients to form the estimate $\hat{s}'$ from the quantized sensor observations received through the non-ideal channels (c.f., (7)). That is why we can view $\hat{s}$ and $\hat{s}'$ in (7) as optimal and pseudo-optimal LMMSE estimates, respectively. We can see from (11) that the variable γ_i naturally factors out from α_i for all i. The advantage of defining the auxiliary variables γ_i's and β_i's in the context of the joint quantization and power allocation will become clear in the subsequent development. There, we shall see that these variables enable us to formulate the problem as a convex optimization problem and based on the resulting solution we propose a block-coordinate descent type algorithm which can achieve monotonic decrease in the distortion.

In the characterization of D, we have assumed error free transmission of the quantized sensor observations to the FC—that is, we assumed ideal communication channels. However, in practice, the channels from the sensors to the FC may subject to fading and additive noise causing bit errors which further increase the distortion. The effect of channel errors on the distortion D_{t} is accounted for by the term D_{c}. We assume that the receiver noise is Gaussian distributed. Herein we also assume that the channel does not change during the transmission of the quantization bits corresponding to each observation cycle—that is, during the transmission of L_i bits—and the channel gain is perfectly known at the FC. Moreover, we assume that the sensor i quantizes the observations to one of the 2^{L_i} quantization levels, encodes the corresponding quantization index using L_i-bit natural binary code, and subsequently transmits these bits to the FC using BPSK modulation scheme. Suppose sensor i expends $\omega_{i,\kappa} \geq 0$ fraction of its transmit power P_i to transmit the κth quantization bit, $\kappa = 1, \ldots, L_i$, then the corresponding bit-error probability is given by $\varepsilon_{i,\kappa} = \mathcal{Q}\left(\sqrt{\zeta_i \omega_{i,\kappa} P_i}\right)$, where $\zeta_i = g_i^2/\sigma_{w_i}^2$ defines the so-called channel SNR and $\sum_{\kappa=1}^{L_i} \omega_{i,\kappa} = 1$. For the contribution of channel bit errors to the distortion, we have the following proposition.

***Proposition** 1:* Under the preceding assumptions, the contribution of channel bit errors to the distortion, characterized by $D_{\mathrm{c}} = \mathbb{E}\left[(\hat{s}' - \hat{s})^2\right]$, can be upper bounded as follows:

$$D_{\mathrm{c}} \leq 3F \sum_{i=1}^{N} \sum_{\kappa=1}^{L_i} \alpha_i^2 L_i 4^{-\kappa} e^{-\zeta_i \omega_{i,\kappa} P_i/2} := D_{\mathrm{c}}^{(\mathrm{opb})}, \tag{17}$$

where $F = 4NW^2/3$.

Proof: See Appendix A. □

Remark 2: From (17) we can see that the contribution of the errors in quantization bits $\kappa = 1, \ldots, L_i$, for all i, to the distortion D_{c} is according to the importance of the bits: The distortion decreases exponentially with the bit-index κ such that the most significant bit (MSB) that is, $\kappa = 1$ when received incorrectly incurs the highest penalty whereas the least significant bit (LSB) that is, $\kappa = L_i$ the smallest penalty in terms of MSE. Therefore, to minimize the distortion, there is a room to optimize transmit power across the sensors as well as along the quantization bits of each sensor.

With (16) and (17) the distortion $\tilde{D}_{\mathrm{t}}$ becomes

$$\begin{aligned}\tilde{D}_{\mathrm{t}} = D + D_{\mathrm{c}} &\leq \sigma_s^2 - \sum_{i=1}^{N} \alpha_i \sigma_s \sigma_{s_i} \rho_{s,i} + 3F \\ &\qquad \sum_{i=1}^{N} \sum_{\kappa=1}^{L_i} \alpha_i^2 L_i 4^{-\kappa} e^{-\zeta_i \omega_{i,\kappa} P_i/2} \\ &= \sigma_s^2 - \mathbf{c}^T \left(\tilde{\mathbf{C}}_{\mathbf{s}} + \mathbf{C}_\gamma^{-1}\right)^{-1} \mathbf{c} + F \mathbf{c}^T \left(\tilde{\mathbf{C}}_{\mathbf{s}} + \mathbf{C}_\gamma^{-1}\right)^{-1} \\ &\qquad \tilde{\mathbf{U}} \left(\tilde{\mathbf{C}}_{\mathbf{s}} + \mathbf{C}_\gamma^{-1}\right)^{-1} \mathbf{c} := \tilde{D}_{\mathrm{t}}^{(\mathrm{opb})},\end{aligned} \tag{18}$$

where $\tilde{\mathbf{U}} = \mathrm{diag}\left(\sum_{\kappa=1}^{L_1} L_1 4^{-\kappa} e^{-\zeta_1 \omega_{1,\kappa} P_1/2}, \ldots, \sum_{\kappa=1}^{L_N} L_N 4^{-\kappa} e^{-\zeta_N \omega_{N,\kappa} P_N/2}\right)$ and $\tilde{D}_{\mathrm{t}}^{(\mathrm{opb})}$, similarly $D_{\mathrm{c}}^{(\mathrm{opb})}$ given in (17), signifies that the power P_i of the sensor i is assigned to the quantization bits L_i according to their significance—where *opb* in the superscript signifies the optimal power allocation to the bits.

In order to simplify the optimization process and for the sake of mathematical tractability, first we assume that the total transmit power P_i of sensor i, for all i, is divided equally among the quantization bits of the sensor that is, $\omega_{i,\kappa} = 1/L_i$ for all κ and use the resulting distortion function to derive an algorithm for the AJQPA by solving a relaxed problem. Afterwards, we allocate power of each sensor to its quantization bits according to their importance.

***Proposition** 2:* Under the assumption of uniform power allocation to the bits, the channel distortion term D_{c} can be upper bounded as follows:

$$D_{\mathrm{c}} \leq F \sum_{i=1}^{N} \alpha_i^2 L_i e^{-\zeta_i P_i / 2L_i} := D_{\mathrm{c}}^{(\mathrm{upb})}. \tag{19}$$

Proof: See Appendix A □

Now with (19) the distortion $\tilde{D}_{\mathrm{t}}$ can be written as follows:

$$\begin{aligned}\tilde{D}_{\mathrm{t}} &\leq \sigma_s^2 - \sum_{i=1}^{N} \alpha_i \sigma_s \sigma_{s_i} \rho_{s,i} + F \sum_{i=1}^{N} \alpha_i^2 L_i e^{-\zeta_i P_i/2L_i} \\ &= \sigma_s^2 - \mathbf{c}^T \left(\tilde{\mathbf{C}}_{\mathbf{s}} + \mathbf{C}_\gamma^{-1}\right)^{-1} \mathbf{c} + F \mathbf{c}^T \left(\tilde{\mathbf{C}}_{\mathbf{s}} + \mathbf{C}_\gamma^{-1}\right)^{-1} \\ &\qquad \mathbf{U} \left(\tilde{\mathbf{C}}_{\mathbf{s}} + \mathbf{C}_\gamma^{-1}\right)^{-1} \mathbf{c} := \tilde{D}_{\mathrm{t}}^{(\mathrm{upb})},\end{aligned} \tag{20}$$

where $\mathbf{U} = \mathrm{diag}\left(L_1 e^{-\zeta_1 P_1/2L_1}, \ldots, L_N e^{-\zeta_N P_N/2L_N}\right)$. Note that the distortions $D_{\mathrm{c}}^{(\mathrm{upb})}$ and $\tilde{D}_{\mathrm{t}}^{(\mathrm{upb})}$ signify that the total power P_i of each sensor i is distributed uniformly among its quantization bits L_i—where *upb* in the superscript signifies the uniform power allocation to the bits.

Now we base the design of the quantization and power allocation scheme on an optimization problem where we minimize the distortion $\tilde{D}_{\mathrm{t}}^{(\mathrm{upb})}$ subject to a constraint on the sum power over all sensors as follows:

$$\begin{aligned}&\underset{L_i, P_i, \forall i}{\mathrm{minimize}}\ \tilde{D}_{\mathrm{t}}^{(\mathrm{upb})} \\ &\text{subject to } \sum_{i=1}^{N} P_i \leq P_{\mathrm{t}},\, L_i \in \mathbb{Z}_+,\, P_i \in \mathbb{R}_+, \quad \forall i.\end{aligned} \tag{21}$$

Since L_i is a positive integer variable (i.e., $L_i \in \mathbb{Z}_+$), the optimization over the quantization variable $L_i \in \mathbb{Z}_+$ and the transmit power $P_i \in \mathbb{R}_+$ constitutes a mixed integer nonlinear programming (MINLP) problem whose computational complexity increases exponentially with the problem size [27]. In order to simplify the problem, we relax the integrality of L_i and allow it to take positive real values (i.e., $L_i \in \mathbb{R}_+$). With this relaxation assumption we can consider the following problem.

$$\begin{aligned}&\underset{\gamma_i, P_i, \forall i}{\text{minimize}}\ \tilde{D}_{\rm t}^{(\rm upb)}\\&\text{subject to}\ \sum_{i=1}^{N} P_i \leq P_{\rm t},\ 0 \leq \gamma_i \leq \frac{1}{\sigma_i^2},\ P_i \geq 0, \quad \forall i,\end{aligned} \tag{22}$$

where the lower bound and the upper bound on γ_i come from (12) for $L_i = 0$ and $L_i = \infty$, respectively. Note that we choose optimization over γ_i's because it is simpler to do than over L_i's. As the γ_i and L_i have one-to-one relationship via (14), so once we know γ_i we can find L_i. Having found $L_i \in \mathbb{R}_+$, we can convert it to its original domain that is, $\mathbb{Z}_+$. In Section III-B, we shall say more about the method to migrate from the relaxed solution to the integer solution.

In order to solve the problem (22) and to derive an algorithm for the joint quantization and power allocation, we adopt the approach where we first optimize the bits for fixed power allocation to the sensors; then we optimize power for given bits. In this way, the problem is decoupled into two subproblems and the solution of these is outlined in Section III-A1 and Section III-A2, which are coming next.

1) Optimizing the quantization bits: For given P_i's, the optimization problem (22) becomes

$$\begin{aligned}&\underset{\gamma_i, \forall i}{\text{minimize}}\ \tilde{D}_{\rm t}^{(\rm upb)}\\&\text{subject to}\ 0 \leq \gamma_i \leq \frac{1}{\sigma_i^2}, \quad \forall i.\end{aligned} \tag{23}$$

Due to the spatial correlations, the objective function is a nonlinear function of the optimization variables and a closed-form solution to the problem is hard to find. To this end, in what follows, we outline an iterative descent-type algorithm which is guaranteed to converge to the Karush-Kuhn-Tucker (KKT) point of the underlying optimization problem.

Let $\gamma = [\gamma_1, \ldots, \gamma_N]^T$ be the vector of γ_i's and $\mathbf{d}\big(\gamma^{(\kappa)}\big) = \nabla_\gamma \tilde{D}_{\rm t}^{(\rm upb)}\big|_{\gamma=\gamma^{(\kappa)}}$ be the gradient of $\tilde{D}_{\rm t}^{(\rm upb)}$ with respect to γ evaluated at $\gamma^{(\kappa)}$, where κ is an iteration index. For simplicity, we denote $\mathbf{d}\big(\gamma^{(\kappa)}\big)$ by $\mathbf{d}^{(\kappa)}$. The ith component of the vector $\mathbf{d}$ is given by

$$\begin{aligned}\mathbf{d}_i &= \frac{\partial \tilde{D}_{\rm t}^{(\rm upb)}}{\partial \gamma_i}\\&= \mathbf{c}^T \left(\tilde{\mathbf{C}}_{\mathbf{s}} + \mathbf{C}_\gamma^{-1}\right)^{-1} \Bigg[\frac{-1}{\gamma_i^2}\mathbf{J}_i + \frac{F}{\gamma_i^2}\mathbf{J}_i \left(\tilde{\mathbf{C}}_{\mathbf{s}} + \mathbf{C}_\gamma^{-1}\right)^{-1}\mathbf{U}\\&\qquad + F\frac{\partial \mathbf{U}_{[i,i]}}{\partial \gamma_i}\mathbf{J}_i + \frac{F}{\gamma_i^2}\mathbf{U}\left(\tilde{\mathbf{C}}_{\mathbf{s}} + \mathbf{C}_\gamma^{-1}\right)^{-1}\mathbf{J}_i\Bigg]\\&\qquad\qquad \left(\tilde{\mathbf{C}}_{\mathbf{s}} + \mathbf{C}_\gamma^{-1}\right)^{-1}\mathbf{c},\end{aligned} \tag{24}$$

where $\mathbf{J}_i$ is a diagonal matrix with unity at (i,i)th place and all other elements equal to zero, and $\frac{\partial \mathbf{U}_{[i,i]}}{\partial \gamma_i} = \frac{1+\zeta_i P_i/2L_i}{2\log(2)\gamma_i\left(1-\gamma_i\sigma_i^2\right)} e^{-\zeta_i P_i/2L_i}$ with L_i given in (14) as a function of γ_i. Note that $\mathbf{U}_{[i,i]}$ denotes (i,i)th element of the matrix $\mathbf{U}$.

Let $\mathcal{P} = \{\gamma_i | 0 \leq \gamma_i \leq 1/\sigma_i^2,\ i = 1, \ldots, N\}$ define the feasible region of the problem (23). The basic idea of the descent-type algorithm is the following: Given $\gamma^{(\kappa)} \in \mathcal{P}$ in iteration κ, find $\gamma^{(\kappa+1)} \in \mathcal{P}$ such that

$$\tilde{D}_{\rm t}^{(\rm upb)}\left(\gamma^{(\kappa+1)}\right) < \tilde{D}_{\rm t}^{(\rm upb)}\left(\gamma^{(\kappa)}\right) \tag{25}$$

by taking a step in the descent direction $\mathbf{\Theta}^{(\kappa)}$; that is, $\gamma^{(\kappa+1)} = \gamma^{(\kappa)} + \tau^{(\kappa)}\mathbf{\Theta}^{(\kappa)}$, where τ is a positive step-length parameter. At $\gamma^{(\kappa)}$, the $\mathbf{\Theta}^{(\kappa)}$ is a descent direction provided $\left(\mathbf{d}^{(\kappa)}\right)^T\mathbf{\Theta}^{(\kappa)} < 0$ holds. To find the descent direction at $\gamma^{(\kappa)}$, assume $\mathbf{\Theta}^{(\kappa)} = \bar{\gamma}^{(\kappa)} - \gamma^{(\kappa)}$ and solve for $\bar{\gamma}^{(\kappa)}$ as follows:

$$\bar{\gamma}^{(\kappa)} = \arg\min_{\gamma \in \mathcal{P}} \left(\mathbf{d}^{(\kappa)}\right)^T\left(\gamma - \gamma^{(\kappa)}\right), \tag{26}$$

where the optimization problem is a linear programming (LP) problem with polyhedron (defined by $\mathcal{P}$) as a feasible region, which is defined by simple bound constraints. The LP problem can be efficiently solved by numerical methods for example, the interior-point method [26]. For $\tau^{(\kappa)} \in [0,1]$ and for $\gamma^{(\kappa)} \in \mathcal{P}$, the $\gamma^{(\kappa+1)} = \gamma^{(\kappa)} + \tau^{(\kappa)}\big(\bar{\gamma}^{(\kappa)} - \gamma^{(\kappa)}\big)$ being a convex combination of two points, $\bar{\gamma}^{(\kappa)}$ and $\gamma^{(\kappa)}$, in $\mathcal{P}$ (a convex set), we have $\gamma^{(\kappa+1)} \in \mathcal{P}$.

Now optimize γ_i's as follows:

a) Initialize $\gamma^{(0)} \in \mathcal{P}$ and set $\kappa = 0$
b) Find $\mathbf{\Theta}^{(\kappa)}$ such that $\left(\mathbf{d}^{(\kappa)}\right)^T\mathbf{\Theta}^{(\kappa)} < 0$ holds by solving (26), otherwise stop.
c) Find step length $\tau^{(\kappa)}$ such that (25) holds.

The iterative procedure generates points $\gamma^{(\kappa)}$'s which are feasible (i.e., $\gamma^{(\kappa)} \in \mathcal{P}$) and decrease $\tilde{D}_{\rm t}^{(\rm upb)}$. It is fairly simple to show that the sequence $\left\{\gamma^{(\kappa)}\right\}$ converges to a stationary point that is, to the first-order KKT point of the underlying optimization problem (see Prop. 2.2.1 in [25]).

Note that the numerically search for γ_i's as outlined requires a number of matrix inversions and for a sensor network with large N it may become computationally expensive. In that case, to reduce the complexity, we adopt the following successive approximation approach to optimize the γ_i's. Noting that $\alpha_i = \beta_i\gamma_i$, we can write

$$\tilde{D}_{\rm t}^{(\rm upb)} = \sigma_s^2 - \sum_{i=1}^{N}\gamma_i\beta_i\sigma_s\sigma_{s_i}\rho_{s,i} + F\sum_{i=1}^{N}\gamma_i^2\beta_i^2\mathbf{U}_{[i,i]}. \tag{27}$$

To optimize γ_i's, we assume that β_i's are given. Then we update the values of β_i's using the values of γ_i's from the previous iteration. Using this idea of successive approximation, the $\tilde{D}_{\rm t}^{(\rm upb)}$ decouples along the sensors and we can solve for γ_i as

$$\gamma_i = \arg\min_{0 \leq \gamma_i \leq \frac{1}{\sigma_i^2}} f(\gamma_i) = -\gamma_i\beta_i\sigma_s\sigma_{s_i}\rho_{s,i} + F\gamma_i^2\beta_i^2\mathbf{U}_{[i,i]}$$

for which the optimality conditions are:

$$\frac{\partial f(\gamma_i)}{\gamma_i} \geq 0, \quad \text{if } \gamma_i = 0, \tag{28}$$

$$\frac{\partial f(\gamma_i)}{\gamma_i} = 0, \quad \text{if } 0 < \gamma_i < \frac{1}{\sigma_i^2}, \tag{29}$$

$$\frac{\partial f(\gamma_i)}{\gamma_i} \leq 0, \quad \text{if } \gamma_i = \frac{1}{\sigma_i^2}. \tag{30}$$

As $\gamma_i = 1/\sigma_i^2$ means $L_i = \infty$ from (14), which is not possible. Therefore, from (28) and (29) find $0 \leq \gamma_i < 1/\sigma_i^2$ by solving

$$2\gamma_i \mathbf{U}_{[i,i]} + \gamma_i^2 \frac{\partial \mathbf{U}_{[i,i]}}{\partial \gamma_i} \geq \frac{\sigma_s \sigma_{s_i} \rho_{s,i}}{F\beta_i}, \quad \forall i. \tag{31}$$

For given P_i's and β_i's, as shown in Appendix B, the function $\tilde{D}_{\mathrm{t}}^{(\mathrm{upb})}$ is convex with respect to γ_i's. Therefore the optimization of γ_i's as given in (31) decreases the distortion.

2) Optimizing the transmit powers: For given L_i's, the optimization problem (22) becomes

$$\begin{aligned} &\underset{P_i,\forall i}{\text{minimize}}\ \tilde{D}_{\mathrm{t}}^{(\mathrm{upb})} \\ &\text{subject to} \sum_{i=1}^{N} P_i \leq P_{\mathrm{t}},\ P_i \geq 0, \quad \forall i, \end{aligned} \tag{32}$$

which we solve using the method of Lagrange multipliers. We define the Lagrangian Λ associated with the problem (32) as

$$\Lambda\left(\lambda, \eta_i, P_i; \forall i\right) = \tilde{D}_{\mathrm{t}}^{(\mathrm{upb})} + \sum_{i=1}^{N} P_i(\lambda - \eta_i) - \lambda P_{\mathrm{t}}, \tag{33}$$

where λ and η_i are Lagrange multipliers. The KKT optimality conditions are given as follows:

$$\frac{\partial \Lambda}{\partial P_i} = \frac{\partial \tilde{D}_{\mathrm{t}}^{(\mathrm{upb})}}{\partial P_i} + \lambda - \eta_i = 0, \quad \forall i; \tag{34}$$

$$\lambda\left(\sum_{i=1}^{N} P_i - P_{\mathrm{t}}\right) = 0,\ \lambda \geq 0,\ \sum_{i=1}^{N} P_i \leq P_{\mathrm{t}}; \tag{35}$$

$$\eta_i P_i = 0,\ \eta_i \geq 0,\ P_i \geq 0, \quad \forall i. \tag{36}$$

In order to assign power to the sensors, solving the KKT conditions in (34) and (36) gives

$$P_i = \frac{2L_i}{\zeta_i} \log^{+}\left(\frac{F\alpha_i^2 \zeta_i}{2\lambda L_i}\right), \quad \text{for } i = 1, \ldots, N, \tag{37}$$

where $\log^{+}(x) = \max\{\log(x), 0\}$ and $\log$ denotes the logarithm to the base e. Eq. (37) shows that the powers allotted to the sensors depend on the correlation values and the observations quality (via α_i's), the quantization bits, and the channel SNRs. Moreover, depending on the values of these parameters some of the sensors with $L_i > 0$ (sensors with $L_i = 0$ always have $P_i = 0$) may get zero power that is, would be switched off altogether. In fact, for a sensor i to get $P_i > 0$ that is, to be active, following conditions must hold: $L_i > 0$ and $F\alpha_i^2\zeta_i/2\lambda L_i > 1$.

***Proposition** 3:* The objective function $\tilde{D}_{\mathrm{t}}^{(\mathrm{upb})}$ is a decreasing and jointly convex function of P_i's for given quantization bits.

Proof: See Appendix C. □

The preceding proposition tells us that the objective function is decreasing with increasing power budget. As we are minimizing a decreasing function, the optimal solution for P_i's is always at the sum-power constraint boundary. Consequently the multiplier λ should be determined so that it satisfies the sum-power constraint with equality that is, $\sum_{j\in\mathcal{A}} P_j = P_{\mathrm{t}}$ which gives

$$\lambda = \exp\left(\frac{P_{\mathrm{t}} + \sum_{j\in\mathcal{A}} \frac{2L_j}{\zeta_j} \log \frac{2L_j}{F\alpha_j^2\zeta_j}}{-\sum_{j\in\mathcal{A}} \frac{2L_j}{\zeta_j}}\right), \tag{38}$$

where $\mathcal{A} := \left\{i | L_i > 0 \wedge F\alpha_i^2\zeta_i/2\lambda L_i > 1,\, i = 1, \ldots, N\right\}$ defines the set of active sensors.

The Prop. 3 tell us that the objective function is jointly convex over all P_i's for given quantization bits. Moreover, the power constraint is linear. Thus the power allocation problem for given quantization bits is convex. Therefore, the power allocation solution given in (37)–(38) is optimal.

Substituting (38) in (37) and using the log–sum inequality [28] we can write

$$P_i \geq \frac{2L_i}{\zeta_i}\left(\frac{P_{\mathrm{t}}}{\sum_{j\in\mathcal{A}} \frac{2L_j}{\zeta_j}} + \log \frac{\alpha_i^2 \zeta_i}{L_i} \frac{\sum_{j\in\mathcal{A}} \frac{L_j}{\zeta_j}}{\sum_{j\in\mathcal{A}} \alpha_j^2}\right). \tag{39}$$

For $P_{\mathrm{t}} \to \infty$, the logarithmic term in (39) is potentially negligible compared to the other term and thus we can approximate P_i as

$$\lim_{P_{\mathrm{t}}\to\infty} P_i \approx \frac{L_i}{\zeta_i} \frac{P_{\mathrm{t}}}{\sum_{j\in\mathcal{A}} L_j/\zeta_j}, \tag{40}$$

which shows that at relatively high P_{t}, the P_i's can be determined by the values of L_i/ζ_i for all i.

3) Summary: Now based on the development in Section III-A1 and Section III-A2, a block-coordinate descent algorithm can be proposed to iteratively optimize the quantization bits and the transmit powers as follows.

I) Initialize the parameters L_i and P_i, and calculate γ_i and α_i for $i = 1, \ldots, N$.
II) Determine power allocation P_i for $i = 1, \ldots, N$ as outlined in Section III-A2.
III) Determine quantization bits L_i for $i = 1, \ldots, N$: find γ_i's as outlined in Section III-A1 and calculate L_i's from (14). Update α_i's from (15).
IV) Repeat steps (II) and (III) until there is no appreciable decrease in the objective function.

We call this as Algorithm 1. The algorithm iteratively optimizes L_i's and P_i's. The algorithm decreases the objective function $\tilde{D}_{\mathrm{t}}^{(\mathrm{upb})}$ in each step—in step (II) for given γ_i's and in step (III) for given P_i's. Therefore the given algorithm can achieve a monotonic decrease in the objective function from one iteration to the other.

B. Migration from Continuous Solution to Integer Solution

In the preceding section, we outlined solution for the continuous relaxation of the MINLP problem. One way to obtain the integer solution is by rounding L_i, for all i, to the nearest integer. However, this constitutes a naive approach. Herein, to obtain the integer solution from the continuous relaxation, we propose a greedy-heuristic procedure outlined in the following, which gives distortion performance better than the simple rounding scheme as we shall see in Section IV.

I) Sort L_i's in an ascending order and construct the corresponding index-set denoted by $\mathcal{L}$.
II) For $i = 1, \ldots, N$, do
III) Set either $L_{\mathcal{L}(i)} = \text{ceil}\left(L_{\mathcal{L}(i)}\right)$ or $L_{\mathcal{L}(i)} = \text{floor}\left(L_{\mathcal{L}(i)}\right)$ depending on which gives the minimum distortion, and construct the index-set $\mathcal{I}$ as $\mathcal{I}(i) = \mathcal{L}(i)$.
IV) Use Algorithm 1 to reallocate power P_k to the sensors $k = 1, \ldots, N$ and bits L_j to the sensors $j \in \mathcal{L} \setminus \mathcal{I}$, where $\mathcal{L} \setminus \mathcal{I}$ means all elements of $\mathcal{L}$ which are not in $\mathcal{I}$.
V) Sort L_j for $j \in \mathcal{L} \setminus \mathcal{I}$ in an ascending order and accordingly update the index-set $\mathcal{L}$.

We call this procedure as Algorithm 2. The given algorithm successively converts the real-valued quantization bits L_i's (from the solution of the relaxed problem) to the integer values starting from the sensor with the smallest number of bits. After converting the relaxed number of bits of the sensor to the integer domain, the algorithm recalculates the bits of the remaining sensors and reallocates power to all sensors using Algorithm 1. In each step of the algorithm, the underlying hypothesis, in converting the bits from the real domain to the integer domain, is that the sensor which gets the smallest number of quantization bits by the Algorithm 1 will be the sensor that is least effective in the network (among the sensors with non-integer bits) and rounding the bits of this sensor is least likely to affect the distortion.

C. Power Allocation Along the Quantization Bits

The objective function $\tilde{D}_{\text{t}}^{(\text{upb})}$ of the optimization problem (21) assumes that the total transmit power P_i of sensor i is uniformly distributed among its quantization bits L_i; that is, P_i/L_i is expended to transmit each of the L_i bits. This approach is not optimal as the contribution of the bits to the reconstruction distortion is proportional to the importance of the bits—an error in the MSB gives the highest increase whereas an error in the LSB the smallest increase to the distortion. Nevertheless, the uniform power allocation among the bits has greatly simplified the optimization problem by enabling us to formulate and solve the continuous relaxation counterpart. Once the proposed algorithm for the relaxed problem has converged and we have converted the real-valued L_i's to the integer domain, then we can allocate the resultant power P_i of sensor i to its quantization bits $\kappa = 1, \ldots, L_i$ according to their significance. To this end, we consider the following problem:

$$\begin{aligned} &\underset{\omega_{i,\kappa},\,\forall i,\kappa}{\text{minimize}}\ D_{\text{c}}^{(\text{opb})} \\ &\text{subject to} \sum_{\kappa=1}^{L_i} \omega_{i,\kappa} = 1,\, \forall i,\, \omega_{i,\kappa} \geq 0, \quad \forall i, \kappa, \end{aligned} \tag{41}$$

where $\omega_{i,\kappa}$ denotes the fraction of P_i used to transmit the quantization bit $\kappa \in \{1, \ldots, L_i\}$, and $D_{\text{c}}^{(\text{opb})}$ is given in (17). Here, the objective function and the constraints are separable along the sensors; therefore, the optimization problem (41) can be written as N separate problems for $i = 1, \ldots, N$ as

$$\begin{aligned} &\underset{\omega_{i,\kappa},\,\forall i,\kappa}{\text{minimize}} \sum_{\kappa=1}^{L_i} 4^{-\kappa} e^{-\zeta_i \omega_{i,\kappa} P_i/2} \\ &\text{subject to} \sum_{\kappa=1}^{L_i} \omega_{i,\kappa} = 1,\, \omega_{i,\kappa} \geq 0, \quad \forall \kappa, \end{aligned} \tag{42}$$

which can be solved independently at the sensor node i.

***Proposition** 4:* The objective function of (42) is a decreasing function of $\omega_{i,\kappa}$ for all κ and the optimization problem is jointly convex over $\omega_{i,\kappa}$ for all κ.

Proof: See Appendix D. □

Using the method of Lagrange multipliers and solving the associated KKT conditions, we can prove that the optimal $\omega_{i,j}$ is given by

$$\omega_{i,j} = \frac{2}{\zeta_i P_i} \log^{+}\left(\frac{\zeta_i P_i}{2\varsigma_i\left(4^j\right)}\right), \quad j = 1, \ldots, L_i, \tag{43}$$

where ς_i is a Lagrange multiplier associated with the constraint $\sum_{\kappa=1}^{L_i} \omega_{i,\kappa} = 1$. Eq. (43) shows that the power allocation to the quantization bits is like a waterfilling on the significance of the bits that is, j where $j = 1$ corresponds to the MSB and $j = L_i$ to the LSB. From (43), we can see that $\omega_{i,j} > 0$ if and only if $\zeta_i P_i/2(4^j)\varsigma_i > 1$. Therefore, we can define a function $f(K_i) = \zeta_i P_i/2(4^{K_i})\varsigma_i$ such that $f(K_i) > 1$ for $\kappa = 1, \ldots, K_i$ and $f(K_i) \leq 1$ for $\kappa = K_i + 1, \ldots, L_i$. The multiplier ς_i can be given by

$$\varsigma_i = \exp\left(\frac{1 + \sum_{\kappa=1}^{K_i} \frac{2}{\zeta_i P_i} \log \frac{2(4^\kappa)}{\zeta_i P_i}}{-\frac{2K_i}{\zeta_i P_i}}\right). \tag{44}$$

By substituting (44) in (43), we can write $\omega_{i,j}$ as follows:

$$\begin{aligned} \omega_{i,j} &= \frac{2}{\zeta_i P_i}\left(\frac{1 + \sum_{\kappa=1}^{K_i} \frac{2}{\zeta_i P_i} \log \frac{24^\kappa}{\zeta_i P_i}}{\frac{2K_i}{\zeta_i P_i}} - \log \frac{2(4^j)}{\zeta_i P_i}\right) \\ &\geq \frac{2}{\zeta_i P_i}\left(\frac{\zeta_i P_i}{2K_i} + \log \frac{2K_i}{\zeta_i P_i \sum_{\kappa=1}^{K_i} 4^{-\kappa}} - \log \frac{2(4^j)}{\zeta_i P_i}\right) \\ &\geq \frac{2}{\zeta_i P_i}\left(\frac{\zeta_i P_i}{2K_i} + \log \frac{6K_i}{\zeta_i P_i} - \log \frac{2(4^j)}{\zeta_i P_i}\right), \\ &= \frac{1}{K_i} + \frac{2}{\zeta_i P_i} \log \frac{3K_i}{4^j}, \end{aligned} \tag{45}$$

where the first inequality follows from the log–sum inequality and the second inequality is obtained by the approximation $\sum_{\kappa=1}^{K_i} 4^{-\kappa} = \left(1 - 4^{-K_i}\right)/3 \leq 1/3$. From (45), it is interesting to note that

$$\lim_{P_i \to \infty} \omega_{i,j} \approx \frac{1}{K_i}, \quad j = 1, \ldots, K_i, \tag{46}$$

which coincides with the classical waterfilling strategy—provided the available power is high enough then it is equally divided among the the bits.

D. Exhaustive Search Based Solution

For the quantization and power allocation design in Section III-A through Section III-C, assuming that the total transmit power of each sensor is uniformly distributed among its quantization bits, we solved the (continuous) relaxed problem; afterwards, we converted the continuous solution to the integer solution; and then based on this integer solution we allotted power to the individual quantization bits. This approach simplified the optimization problem because it enabled us to relax the integrality constraint on the quantization bits. Alternatively, noting that $P_{i,\kappa} = \omega_{i,\kappa} P_i$ is the power expended to transmit the κth quantization bit of sensor i, we can consider the following optimization problem:

$$\underset{L_i, P_{i,\kappa}, \forall i,\kappa}{\text{minimize}} \; \tilde{D}_{\rm t}^{(\rm opb)}$$

$$\text{subject to} \sum_{i=1}^{N} \sum_{\kappa=1}^{L_i} P_{i,\kappa} \le P_{\rm t}, \quad L_i \in \mathbb{Z}_+, \; P_{i,\kappa} \in \mathbb{R}_+, \quad \forall i, \kappa.$$

For given L_i's, using the Lagrange multiplier method, the optimal $P_{i,j}$ for all i and j can be given by

$$P_{i,j} = \frac{2}{\zeta_i} \log^+ \left(\frac{3F\alpha_i^2 L_i \zeta_i}{2\lambda(4^j)} \right), \tag{47}$$

where λ is a Lagrange multiplier associated with the sum-power constraint. The multiplier can be given by

$$\lambda = \exp \left(\frac{P_{\rm t} + \sum_{i,l \in \mathcal{A}} \frac{2}{\zeta_i} \log \frac{2(4^l)}{3F\alpha^2 L_i \zeta_i}}{-\sum_{i,l \in \mathcal{A}} \frac{2}{\zeta_i}} \right), \tag{48}$$

where $\mathcal{A} := \{i, j | 3F\alpha_i^2 L_i \zeta_i / 2\lambda(4^j) > 0, \, j = 1, \ldots, L_i, i = 1, \ldots, N\}$. Let the optimal $P_{i,j}$'s be denoted by $P_{i,j}^*$'s then the optimal L_i's can be obtained by exhaustive search as a solution to the following:

$$L_i^*, \forall i = \arg \min_{L_i \in \{0, \ldots, L_{\max}\}, \forall i} \tilde{D}_{\rm t}^{(\rm opb)} \left(L_i, P_{i,j}^*(L_i) \right), \tag{49}$$

where $L_{\max} \in \mathbb{Z}_{++}$ is some reasonably large integer with $\mathbb{Z}_{++}$ being the set of all strictly positive integers. Note that the computational cost of such a search exponentially increases with N as the search domain expands as $(1 + L_{\max})^N - 1$.

IV. Numerical and Simulation Examples

Through numerical and simulation examples, this section substantiates the analytical findings and illustrates the effectiveness of the quantization and power allocation scheme. To this purpose, we consider an elementary sensor network with $N = 3$, unless stated otherwise. Note that for the sake of illustration we have selected a small network size. Nevertheless, the simulation of a network of any size can similarly be done, as we shall see later in this section. We assume without any loss of generality that $\theta_2 = 1$, $\sigma_s^2 = \sigma_{s_i}^2 = 1$, $\sigma_{n_i}^2 = 0.01$, $g_i = 1$, and $\sigma_{w_i}^2 = 1$ for all i. For the sake of illustration, we take the following example $(d_{X_1}, d_{X_2}, d_{X_3}) = (-0.1, \, 0, \, 1.5)$ and $(d_{Y_1}, d_{Y_2}, d_{Y_3}) = (0, \, 5, \, 0)$, where (d_{X_i}, d_{Y_i}) gives the position of sensor i in the XY-plane. Note that we can view this example as a realization of random deployment of the sensors. We have taken this example for purely illustrative purpose which in no-way limits the generality of the results. We assume that the source lies at the origin of the XY-plane. Assuming $\theta_1 = 1$, the corresponding spatial correlation values are $(\rho_{s,1}, \rho_{s,2}, \rho_{s,3}) = (0.9048, 0.0067, 0.2231)$ and $(\rho_{1,2}, \rho_{1,3}, \rho_{2,3}) = (0.0067, 0.2019, 0.0054)$. Note that for

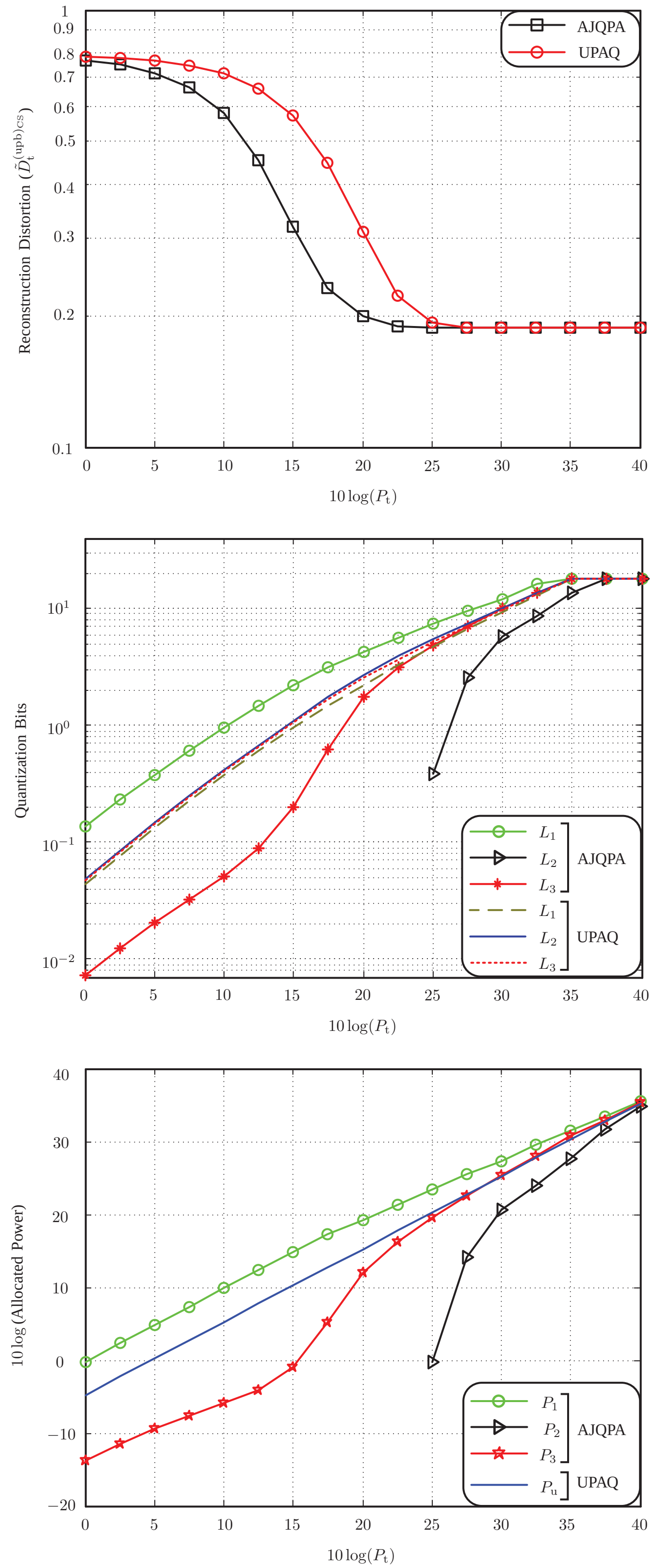

Fig. 2: Comparison of the quantization and power allocation schemes AJQPA and UPAQ .

the given spatial correlation values and the observation noise variances, the lower bound MSE is $D_0 = 0.1875$. The total estimation distortion cannot be below this value no matter how finely we quantize the sensor observations and how large the transmit power may become. Unless stated otherwise, all numerical and simulation examples in this work are based on this system setup. In the figures $\log(.) = \log_{10}(.)$. In the sequel the quantization and power allocation design presented in Section III-A through Section III-C is referred as AJQPA-Ia and the exhaustive search based design in Section III-D as AJQPA-Ib.

First we compare the AJQPA-Ia design (under Algorithm 1) with a uniform power-allocation based quantization (UPAQ) scheme. The UPAQ scheme distributes power uniformly among the sensors that is, $P_i = P_{\mathrm{u}} = P_{\mathrm{t}}/N$ for all i. The quantization bits L_i's for the UPAQ are calculated in the same way as in the AJQPA-Ia design. The results are plotted in Fig. 2, which shows that the AJQPA design outperforms the so-called UPAQ scheme in terms of the reconstruction MSE. This superior performance comes from the fact that, contrary to the UPAQ, the AJQPA design quantizes finely and allocates more power to the sensor(s) having favorable correlation values while some sensors with less favorable correlation values are turned on at higher total power P_{t} or completely switched off. Note that with increasing P_{t}, the achieved distortion approaches the lower bound $D_0 = 0.1875$. Moreover, at high P_{t} the AJQPA design allocates power uniformly among the sensors, and the quantization bits also becomes equal across the sensors which saturates at some finite value—the AJQPA design converges to the UPAQ scheme. The examples in the sequel only consider the AJQPA design without including comparison with the UPAQ scheme. Nevertheless, the AJQPA design always performs better than the so-called UPAQ scheme.

Fig. 3 compares the reconstruction distortions, and the associated quantization bits and power allocations obtained from Algorithm 1 for the solution of continuous relaxation, denoted as continuous solution (CS), and the integer solution (IS) obtained from the CS by Algorithm 2 proposed under the AJQPA-Ia scheme. The figure also plots the quantization bits, the transmit power of the sensors, and the associated distortion under the scheme AJQPA-Ib. In the figure, the quantities concerning the AJQPA-Ib are denoted by the extension "-ES"—signifying the exhaustive search based solution. From the figure, we can observe that, under the AJQPA-Ia, the CS obtained from Algorithm 1 and the IS obtained from Algorithm 2 give distortions $\tilde{D}_{\mathrm{t}}^{(\mathrm{upb})_{\mathrm{CS}}}$ and $\tilde{D}_{\mathrm{t}}^{(\mathrm{upb})_{\mathrm{IS}}}$, respectively, which monotonically decrease with P_{t}. Similarly the distortion $\tilde{D}_{\mathrm{t}}^{(\mathrm{opb})_{\mathrm{IS\text{-}ES}}}$ obtained from the AJQPA-Ib shows a monotonic decrease with P_{t}. Moreover, note that $\tilde{D}_{\mathrm{t}}^{(\mathrm{opb})_{\mathrm{IS\text{-}ES}}} \leq \tilde{D}_{\mathrm{t}}^{(\mathrm{opb})_{\mathrm{IS}}} \leq \tilde{D}_{\mathrm{t}}^{(\mathrm{upb})_{\mathrm{IS}}}$, where $\tilde{D}_{\mathrm{t}}^{(\mathrm{upb})_{\mathrm{IS}}}$ denotes the distortion when power allotted to each sensor is distributed uniformly among its quantization bits whereas $\tilde{D}_{\mathrm{t}}^{(\mathrm{opb})_{\mathrm{IS}}}$ denotes the distortion when power of each sensor is allotted to its bits according to their significance. To this end, Fig. 4 plots the fraction of the power of each sensor allotted to transmit its individual quantization bits. We can observe that the MSB is given the largest share and the LSB the smallest share in the transmit power of the respective sensor.

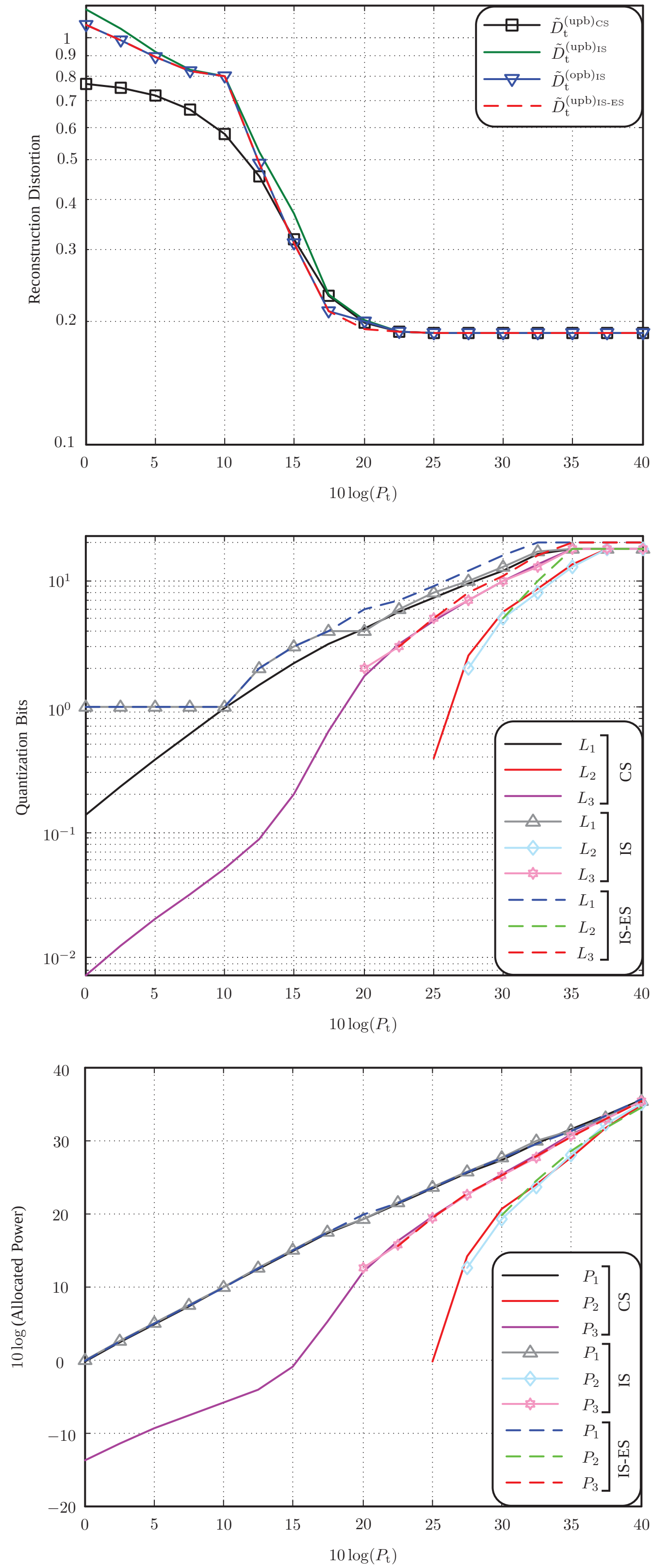

Fig. 3: Comparison of the quantization and power allocation schemes AJQPA-Ia and AJQPA-Ib.

From Fig. 3, we observe that the distortion $\tilde{D}_{\mathrm{t}}^{(\mathrm{opb})_{\mathrm{IS}}}$ achieved by the AJQPA-Ia design and the distortion

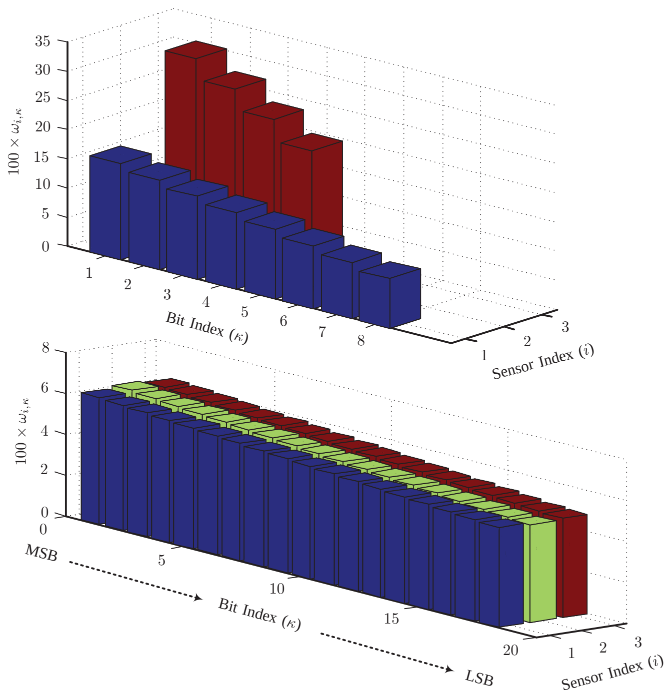

Fig. 4: Distribution of each sensor power P_i to its bits under the AJQPA-Ia scheme. (Top) At $10\log(P_{\rm t}) = 25$, sensors get power $10\log(P_1, P_2, P_3) = (23.7914, \text{Off}, 18.8545)$ and bits $(L_1, L_2, L_3) = (8, 0, 4)$. (Bottom) At $10\log(P_{\rm t}) = 40$, sensors get power $10\log(P_1, P_2, P_3) = (35.5437, 34.8415, 35.2725)$ and bits $(L_1, L_2, L_3) = (18, 18, 18)$.

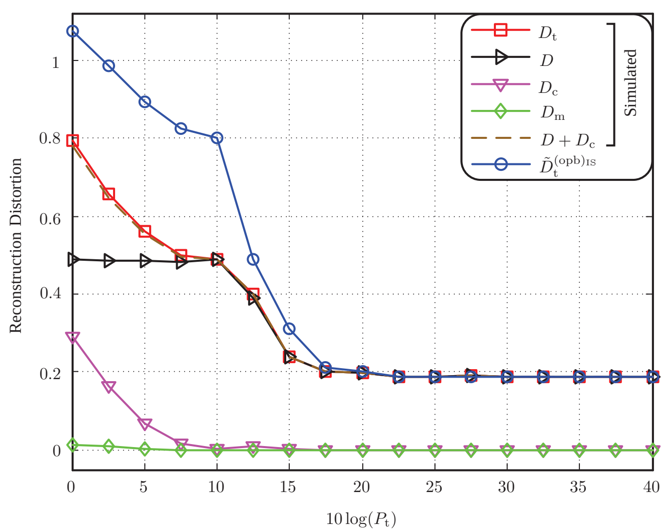

Fig. 5: Distortion performance comparison: theoretical versus simulated.

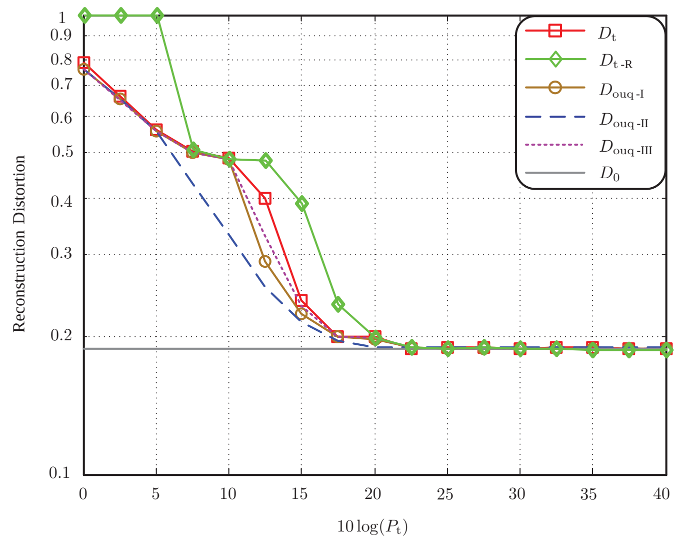

Fig. 6: Distortion performance comparison of the quantization and power allocation schemes AJQPA-Ia and AJQPA-II for $N = 3$.

$\tilde{D}_{\rm t}^{(\rm opb)_{IS\text{-}ES}}$ achieved by the AJQPA-Ib design are quite close to each other. This is because the quantization bits and the power allotted to the sensors by the two designs match quite well. This observation shows that, compared to the exhaustive search based design (i.e., AJQPA-Ib), the design proposed under the AJQPA-Ia scheme works quite well. The simulation examples in the sequel focus on the design AJQPA-Ia, unless stated otherwise.

Examples in Fig. 5 compare the theoretical reconstruction distortion with that obtained from the actual system simulation. We can observe that the distortion term $D_{\rm m}$ is negligibly small and $D_{\rm t} \approx D + D_{\rm c}$. Moreover, the distortion $\tilde{D}_{\rm t}^{(\rm opb)_{IS}}$ is a quite tight upper bound for $D + D_{\rm c}$ for all $P_{\rm t}$ values where any of the sensors quantizes with more than one bit (c.f., Fig. 3). These observations reveal that the pseudo-quantization noise model combined with upper bounding the distortion $D_{\rm t}$ with $\tilde{D}_{\rm t}^{(\rm opb)_{IS}}$ give quite good results.

Next, in Fig. 6 we compare the performance of the quantization and power allocation scheme AJQPA-Ia with that of the scheme in [20] which is referred here as AJQPA-II. In the figure, $D_{\rm t}$ is the total distortion achieved by the AJQPA-Ia scheme; $D_{\text{t-R}}$ is the distortion achieved by a scheme where the CS for the quantization bits from the AJQPA-Ia is converted to the integer domain by simple rounding and then we reallocate transmit power to the rounded bits as is done in AJQPA-Ia; $D_{\text{ouq-I}}$, $D_{\text{ouq-II}}$, and $D_{\text{ouq-III}}$ denote the distortions achieved by the AJQPA-II scheme. The $D_{\text{ouq-I}}$ is obtained by solving the optimization problem in [20] with the $\{L_i\}_{i=1}^N$ obtained from the AJQPA-Ia (c.f., Fig. 3), $D_{\text{ouq-II}}$ is obtained by solving the problem of [20] by exhaustive search over $\{L_i\}_{i=1}^N$, and $D_{\text{ouq-III}}$ is obtained by evaluating the objective function of the problem in [20] for both $\{L_i\}_{i=1}^N$ and $\{P_i\}_{i=1}^N$ provided by the AJQPA-Ia. The D_0 denotes the lower bound distortion. We can observe that $D_0 \leq D_{\text{ouq-II}} \leq D_{\text{ouq-I}} \leq D_{\rm t} \leq D_{\text{t-R}}$. The difference between $D_{\rm t}$ and $D_{\text{t-R}}$ highlights the effectiveness of the Algorithm 2 under the AJQPA-Ia scheme for obtaining the IS from the CS vis-à-vis the simple rounding scheme. Moreover, we can see that $D_{\rm t}$ and $D_{\text{ouq-III}}$ are quite close to both $D_{\text{ouq-I}}$ and $D_{\text{ouq-II}}$ for a wide range of $P_{\rm t}$ values. This illustrates the effectiveness of the AJQPA-Ia scheme when compared with the computational complexity of the AJQPA-II

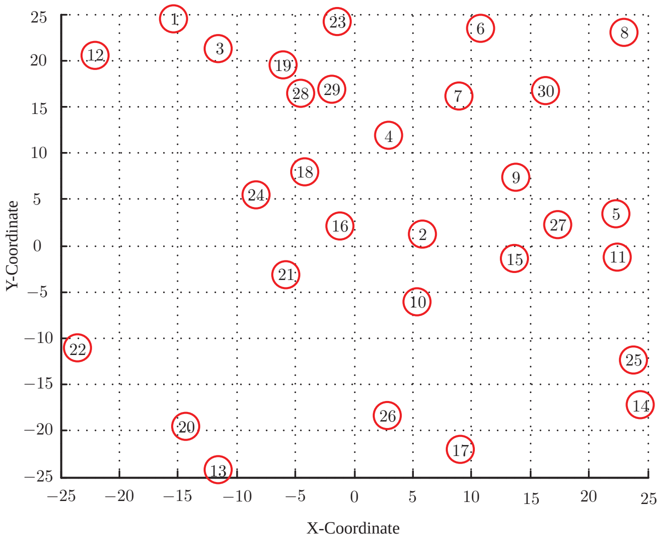

Fig. 7: $N = 30$ sensors deployment layout.

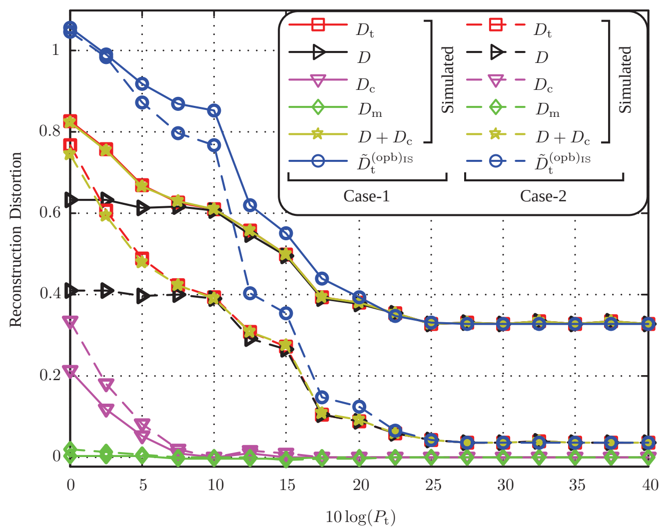

Fig. 8: Distortion performance comparison for $N = 30$: theoretical versus simulated.

scheme, especially in the case of $D_{\text{ouq-II}}$.

The simulation examples thus far are based on a network that consists of three sensors for which the correlation values are fixed. Next, we consider a sensor network comprising thirty sensors that is $N = 30$, which are deployed in a fifty-by-fifty grid as shown in Fig. 7. We assume that the underlying source is located at the center of this grid. Similar to the three-sensor case, we assume $\theta_2 = 1$, $\sigma_s^2 = \sigma_{s_i}^2 = 1$, $\sigma_{n_i}^2 = 0.01$, $g_i = 1$, and $\sigma_{w_i}^2 = 1$ for all i. For this network setup, we consider two cases having different correlation structure: *Case-1* with $\theta_1 = 10$ and *Case-2* with $\theta_1 = 10$, where θ_1 controls how fast the correlation decays with the distance (c.f., (2)). For these cases, Fig. 8 compares the theoretical and simulated distortions. We can see that the distortions, both theoretical and simulated, decay with increasing correlation that is, increasing value of θ_1. Similar to the three-sensor case,

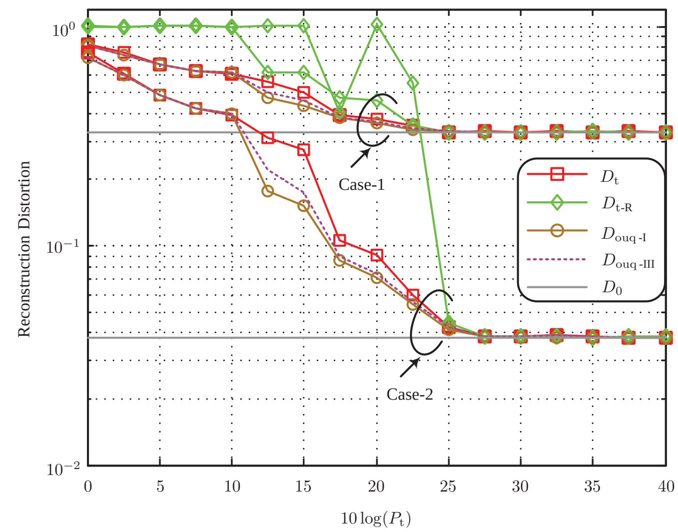

Fig. 9: Distortion performance comparison of the quantization and power allocation schemes AJQPA-Ia and AJQPA-II for $N = 30$.

we observe that $D_t \approx D + D_c$ with D_m being negligibly small. Finally, Fig. 9 gives distortion comparison for the AJQPA-Ia and AJQPA-II schemes, where similar observations can be made like the three-sensor case in Fig. 6. Note that the $D_{\text{ouq-II}}$ is not included in Fig. 9 because it is intractable to obtain for $N = 30$. To obtain $D_{\text{ouq-II}}$, we need to solve the optimization problem of [20] by exhaustive search for $\{L_i\}_{i=1}^{N}$ over $\{0, \ldots, L_{\max}\}$. For example, for $L_{\max} = 9$ we need to solve the problem for 10^{30} times, which is not a computationally manageable task within a reasonable time frame.

V. Conclusion

In this contribution, we have pursued a design to jointly quantize the sensor observations, which are correlated across sensors, and allocate power to transmit the observations to the FC with the goal to reconstruct the source with minimum distortion. Based on the assumption of pseudo-quantization noise model and the quasi-optimal LMMSE estimate, we showed that the quantization and power allocation problem can be solved efficiently. Based on the solution, we proposed a block-coordinate descent type algorithm which iteratively optimizes the quantization and power allocation. Moreover, we showed that in addition to the power allocation across the sensors, there is a room to optimize power allotted to transmit individual quantization bits of each sensor. We illustrated the effectiveness of the proposed designs with a few simple examples. We have seen that sensors having high correlation and low cross-correlation values and better observation quality compared to other sensors quantize their observation with finer resolution and transmit at higher power. It was also shown that the proposed design outperforms a quantization scheme based on the uniform power allocation. Finally, from the simulation examples it appears that the theoretical distortion approximates the simulated value quite well when any of the sensors quantizes with more than one bit. The future work will consider the effect of imperfect knowledge of the correlation, the observation noise, and the communication channel. In this work, we considered the system where FC decodes the received signal before combining them to form the final estimate. However, the future work may consider estimation based on the soft received signals from the sensors.

Appendix A
Channel Bit Errors Contribution to Distortion

The mean-squared error of $\hat{s}'$ with respect to $\hat{s}$ is denoted by D_{c} which can be written as follows:

$$\begin{aligned} D_{\mathrm{c}} &= \mathbb{E}\left[\left(\hat{s}'-\hat{s}\right)^2\right] = \mathbb{E}\left[\left(\sum_{i=1}^{N}\alpha_i\left(m_i'-m_i\right)\right)^2\right] \\ &\leq N\sum_{i=1}^{N}\alpha_i^2\mathbb{E}\left[\left(m_i'-m_i\right)^2\right], \end{aligned} \tag{50}$$

using the Cauchy–Schwarz inequality, where m_i' is the quantized message received through the actual channel and m_i is the quantized message received error free at the FC via ideal channel. The messages m_i and m_i' have L_i bits each and can be written as follows:

$$\begin{aligned} m_i &= \left(\sum_{\kappa=1}^{L_i} b_{i,\kappa}2^{L_i-\kappa} - 2^{L_i-1} + \frac{1}{2}\right)\Delta_i, \\ m_i' &= \left(\sum_{\kappa=1}^{L_i} b_{i,\kappa}'2^{L_i-\kappa} - 2^{L_i-1} + \frac{1}{2}\right)\Delta_i, \end{aligned} \tag{51}$$

where $b_{i,\kappa}, b_{i,\kappa}' \in \{0, 1\}$ for all i and κ. The sensor i transmits the quantization bit κ, $\kappa = 1, \ldots, L_i$, to the FC using BPSK modulation and power $\omega_{i,\kappa}P_i$. The associated bit-error probability is $\varepsilon_{i,\kappa} = \mathcal{Q}\big(\sqrt{\zeta_i\omega_{i,\kappa}P_i}\big)$ where $\zeta_i = g_i^2/\sigma_{w_i}^2$. With this, we have

$$\begin{aligned} \mathbb{E}\left[\left(m_i'-m_i\right)^2\right] &= \Delta_i^2\mathbb{E}\left[\left(\sum_{\kappa=1}^{L_i}2^{L_i-\kappa}\left(b_{i,\kappa}'-b_{i,\kappa}\right)\right)^2\right] \\ &\leq \Delta_i^2 L_i\sum_{\kappa=1}^{L_i}2^{2L_i-2\kappa}\mathbb{E}\left[\left(b_{i,\kappa}'-b_{i,\kappa}\right)^2\right], \end{aligned} \tag{52}$$

where $(b_{i,\kappa}'-b_{i,\kappa})^2$ is a Bernoulli distributed random variable—one with probability $\varepsilon_{i,\kappa}$ and zero with probability $(1-\varepsilon_{i,\kappa})$, and $\mathbb{E}\big[(b_{i,\kappa}'-b_{i,\kappa})^2\big] = \varepsilon_{i,\kappa}$. Therefore (52) can be written as

$$\begin{aligned} \mathbb{E}\left[\left(m_i'-m_i\right)^2\right] &\leq \Delta_i^2 L_i\sum_{\kappa=1}^{L_i}2^{2L_i-2\kappa}\varepsilon_{i,\kappa} \\ &= 4W^2L_i\sum_{\kappa=1}^{L_i}2^{-2\kappa}\varepsilon_{i,\kappa}. \end{aligned} \tag{53}$$

Substituting (53) in (50) we get

$$\begin{aligned} D_{\mathrm{c}} &\leq 4NW^2\sum_{i=1}^{N}\sum_{\kappa=1}^{L_i}\alpha_i^2L_i4^{-\kappa}\varepsilon_{i,\kappa} \\ &\leq 4NW^2\sum_{i=1}^{N}\sum_{\kappa=1}^{L_i}\alpha_i^2L_i4^{-\kappa}e^{-\zeta_i\omega_{i,\kappa}P_i/2}, \end{aligned} \tag{54}$$

where the last inequality follows from $\varepsilon_{i,\kappa} \leq e^{-\zeta_i\omega_{i,\kappa}P_i/2}$. Now, assuming that the total transmit power P_i of sensor i is divided equally among the quantization bits L_i's that is, $\omega_{i,\kappa} = 1/L_i$, we get $\varepsilon_{i,\kappa} = \varepsilon_i = \mathcal{Q}\big(\sqrt{\zeta_iP_i/L_i}\big)$ for all κ. Now, from (53) we can write as follows:

$$\begin{aligned} \mathbb{E}\left[\left(m_i'-m_i\right)^2\right] &\leq 4W^2L_i\varepsilon_i\sum_{\kappa=1}^{L_i}2^{-2\kappa} \\ &= 4W^2L_i\varepsilon_i\frac{1-0.25^{L_i}}{3} \\ &\leq \frac{4W^2}{3}L_i\varepsilon_i \leq \frac{4W^2}{3}L_ie^{-\zeta_iP_i/2L_i}. \end{aligned} \tag{55}$$

Substituting (55) in (50) and assuming $F = 4NW^2/3$ we get

$$D_{\mathrm{c}} \leq F\sum_{i=1}^{N}\alpha_i^2L_ie^{-\zeta_iP_i/2L_i}. \tag{56}$$

Appendix B
Convexity of $\tilde{D}_{\mathrm{t}}^{(\mathrm{upb})}$ over γ_i's

With respect to γ_j's, the second-order partial derivatives of $\tilde{D}_{\mathrm{t}}^{(\mathrm{upb})}$ given in (27) can be written as follows: For all $i \neq j$, we have $\partial^2\tilde{D}_{\mathrm{t}}^{(\mathrm{upb})}/\partial\gamma_j\partial\gamma_i = 0$; moreover, for all i, we have

$$\begin{aligned} \frac{\partial^2\tilde{D}_{\mathrm{t}}^{(\mathrm{upb})}}{\partial\gamma_i^2} &= 2F\beta_i^2\mathbf{U}_{[i,i]} + F^2\beta_i^2\gamma_i\mathbf{U}_{[i,i]}\frac{2\left(1-\gamma_i\sigma_i^2\right)+1}{\left(1-\gamma_i\sigma_i^2\right)} \\ &\quad \left(\frac{\beta_i\zeta_iP_i}{4\log(2)L_i^2\left(1-\gamma_i\sigma_i^2\right)}\right)^2\frac{\partial\mathbf{U}_{[i,i]}}{\partial\gamma_i} \geq 0, \end{aligned}$$

because each of the term in the summation is positive. Thus the Hessian of $\tilde{D}_{\mathrm{t}}^{(\mathrm{upb})}$ with respect to γ_i's is diagonal and each diagonal element is positive meaning that the Hessian is positive semidefinite, which proves that the function is convex over γ_i's for given P_i's and β_i's.

Appendix C
Proof of Prop. 3

The first-order derivative of the function $\tilde{D}_{\mathrm{t}}^{(\mathrm{upb})}$ with respect to P_i is $\partial\tilde{D}_{\mathrm{t}}^{(\mathrm{upb})}/\partial P_i = -F\gamma_i^2\beta_i^2\zeta_i\mathbf{U}_{[i,i]}/2L_i$, which is always negative for any P_i. This tells us that $\tilde{D}_{\mathrm{t}}^{(\mathrm{upb})}$ is a decreasing function of P_i.

The second-order derivatives of $\tilde{D}_{\mathrm{t}}^{(\mathrm{upb})}$ with respect to P_i's are given as follows: $\partial^2\tilde{D}_{\mathrm{t}}^{(\mathrm{upb})}/\partial P_j\partial P_i = 0$ for all $i \neq j$ and $\partial^2\tilde{D}_{\mathrm{t}}^{(\mathrm{upb})}/\partial P_i^2 = F\beta_i^2\gamma_i^2\zeta_i^2\mathbf{U}_{[i,i]}/4L_i^3 \geq 0$ for all i. Thus the Hessian of $\tilde{D}_{\mathrm{t}}^{(\mathrm{upb})}$ with respect to P_i's is diagonal and is positive semidefinite. The positive semidefiniteness of the Hessian means that the function is jointly convex over P_i's for given quantization bits.

Appendix D
Proof of Prop. 4

As the constraints of the given optimization problem are linear, therefore for the problem to be convex it suffices to show that the objective function is convex over the optimization variables $\omega_{i,\kappa}$ for all κ. To this purpose, let $f(\omega_{i,1}, \ldots, \omega_{i,L_i}) =$

$\sum_{\kappa=1}^{L_i} 4^{-\kappa} e^{-\zeta_i \omega_{i,\kappa} P_i/2}$. For $\kappa, \iota \in \{1, \ldots, L_i\}$, we can show that

$$\frac{\partial^2 f}{\partial \omega_{i,\kappa} \partial \omega_{i,\iota}} = 0, \quad \forall \kappa \neq \iota;$$

and

$$\frac{\partial^2 f}{\partial \omega_{i,\kappa}^2} = \left(4^{-\kappa-1}\right) \zeta_i^2 P_i^2 e^{-\zeta_i \omega_{i,\kappa} P_i/2} \geq 0, \quad \forall \kappa,$$

which tells us that the Hessian of $f(\omega_{i,1}, \ldots, \omega_{i,L_i})$ is positive semidefinite and thus proving that the given function is jointly convex over $\omega_{i,\kappa}$'s. Furthermore, note that

$$\frac{\partial f}{\partial \omega_{i,\kappa}} = -\frac{\zeta_i P_i (4^{-\kappa})}{2} e^{-\zeta_i \omega_{i,\kappa} P_i/2},$$

which is always negative for any valid $\omega_{i,\kappa}$ and thus $f(\omega_{i,1}, \ldots, \omega_{i,L_i})$ is a decreasing function over $\omega_{i,\kappa}$.

References

[1] J.-J. Xiao and Z.-Q. Luo, "Multiterminal source-channel communication over an orthogonal multiple-access channel," *IEEE Trans. on Inform. Theory*, vol. 53, no. 9, pp. 3255–3264, Sep. 2007.

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

[3] A. Krasnopeev, J.-J Xiao, and Z.-Q. Luo, "Minimum energy decentralized estimation in a wireless sensor network with correlated sensor noises," *EURASIP Jr. on Wireless Commun. and Networking*, pp. 473–482, 2005.

[4] J. Li and G.Alregib, "Rate-constrained distributed estimation in wireless sensor networks," *IEEE Trans. on Signal Processing*, vol. 55, pp. 1634–1643, May 2007.

[5] X. Luo and G. B. Giannakis, "Energy-constrained optimal quantization for wireless sensor networks," *EURASIP Jr. on Advances in Signal Process.*, vol. 2008, Article ID 462930, 12 pages, 2008.

[6] L. Zhang, X.-D. Zhang, T. Ho, and T. Dikaliotis, "Progressive distributed estimation over noisy channels in wireless sensor networks," in *Proc. of IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP'09)*, pp. 2221–2224, Taipei (Taiwan), Apr. 2009.

[7] H. Li and J. Fang, "Distributed adaptive quantization and estimation for wireless sensor networks," *IEEE Signal Processing Letters*, vol. 14, no. 10, pp. 669–672, Oct. 2007.

[8] Y. Huang and Y. Hua, "Energy planning for progressive estimation in multihop sensor networks," *IEEE Trans. on Signal Processing*, vol. 57, no. 10, pp. 4052–4065, Oct. 2009.

[9] J. Y. Wu, Q. Z. Huang, and T. S. Lee, "Minimal energy decentralized estimation via exploiting the statistical knowledge of sensor noise variance," *IEEE Trans. on Signal Processing*, vol. 56, no. 5, pp. 2171–2176, May 2008.

[10] T. C. Aysal and K. E. Barner "Constrained decentralized estimation over noisy channels for sensor networks," *IEEE Trans. on Signal Processing*, vol. 56, no. 4, pp. 1398–1410, Apr. 2008.

[11] H. Chen and P. K. Varshney "Performance limit for distributed estimation systems with identical one-bit quantizers," *IEEE Trans. on Signal Processing*, Vol. 58, No. 1, pp. 466–471, Jan. 2010.

[12] J. Fang and H. Li "Distributed adaptive quantization for wireless sensor networks: from delta modulation to maximum likelihood," *IEEE Trans. on Signal Processing*, vol. 56, no. 10, pp. 5246–5257, Oct. 2008.

[13] J. Fang and H. Li "Hyperplane-based vector quantization for distributed estimation in wireless sensor networks," *IEEE Trans. on Inform. Theory*, vol. 55, no. 12, pp. 5682–5699, Dec. 2009.

[14] J. Matamoros and C. Antón-Haro, "Bandwidth constraints in wireless sensor-based decentralized estimation schemes for Gaussian channels," in *Proc. of IEEE Global Commun. Conf. (GLOBECOM'08)*, pp. 1–5, Lousiana, New Orleans (USA), Nov.–Dec. 2008.

[15] M. C. Vuran and O. B. Akan, and I. F. Akylidiz, "Spatio-temporal correlation: theory and applications for wireless sensor networks," *Elsevier Jr. on Computer Networks*, vol. 45, no. 3, pp. 245–259, Jun. 2004.

[16] I. bahceci and A. K. Khandani, "Linear estimation of correlated data in wireless sensor networks with optimum power allocation and analog modulation," *IEEE Trans. on Commun.*, vol. 56, no. 7, pp. 1146–1156, Jul. 2008.

[17] F. Davoli, M. Marchese, and M. Mongelli, "A decision theoretic approach to Gaussian sensor networks," in *Proc. of IEEE Int. Conf. on Commun. (ICC'09)*, pp. 1–5, Dresden (Germany), Jun. 2009.

[18] J. Fang and H. Li "Power constrained distributed estimation with correlated sensor data," *IEEE Trans. on Signal Processing*, vol. 57, no. 8, pp. 3292–3297, Aug. 2009.

[19] M. H. Chaudhary and L. Vandendorpe "Adaptive power allocation in wireless sensor networks with spatially correlated data and analog modulation: perfect and imperfect CSI," *EURASIP Jr. on Wireless Commun. and Networking*, vol. 2010, Article ID 817961, 14 pages, 2010.

[20] M. H. Chaudhary and L. Vandendorpe, "Quantization and power allocation in wireless sensor networks with correlated data," in *Proc. of IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP'11)*, pp. 3016–3019, Prague (Czech Republic), May 2011.

[21] B. Widrow and I. Kollar, *Quantization Noise: Roundoff Error in Digital Computation, Signal Processing, Control and Communications,* Cambridge University Press, 2008.

[22] R. M. Gray and D. L. Neuhoff, "Quantization," *IEEE Trans. on Inform. Theory*, vol. 44, no. 6, pp. 2325–2383, Oct. 1998.

[23] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory,* Prentice-Hall, 1993.

[24] S. Murugesan, "Harnessing green IT: principles and practices," *IEEE IT Professional*, vol. 10, no. 1, pp. 24–33, Jan.–Feb. 2008.

[25] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods,* Athena Scientific, Massachusetts, 1997.

[26] S. Boyd and L. Vandenberghe, *Convex Optimization,* Cambridge University Press, 2008.

[27] D. Li and X. Sun, *Nonlinear Integer Programming,* Springer, 2006.

[28] T. M. Cover and J. A. Thomas, *Elements of Information Theory, Second Edition,* John Wiley and Sons, 2006.

Muhammad Hafeez Chaudhary was born in the Punjab region of Pakistan. He attended Govt. College University Lahore, Pakistan, for his pre-engineering studies. He received his B.Sc. Electrical Engineering degree with *honours* from the University of Engineering and Technology Lahore in 2001. He worked at the National Engineering and Scientific Commission Islamabad, Pakistan, from 2001 to 2005 where he was involved in the research and development of sensor systems.

In 2007, he did European Master of Science in Research on Information and Communication Technologies, a double-degree M.S. in electrical engineering jointly organized by the Université catholique de Louvain (UCL), Louvain-La-Neuve, Belgium and the Karlsruhe Institute of Technology, Karlsruhe, Germany with the overall grade *sehr gut* (very good: an *outstanding achievement*). Since 2007, he is affiliated with the Institute of Information and Communication Technologies, Electronics and Applied Mathematics of UCL as a research assistant. His research interests include algorithm and protocol design for information communication and processing in next generation wireless networks.

Luc Vandendorpe (M'93–SM'99–F'06) was born in Mouscron, Belgium, in 1962. He received the Electrical Engineering degree (summa cum laude) and the Ph.D. degree from the Université catholique de Louvain (UCL), Louvain-la-Neuve, Belgium, in 1985 and 1991, respectively.

Since 1985, he has been with the Communications and Remote Sensing Laboratory of UCL, where he first worked in the field of bit rate reduction techniques for video coding. In 1992, he was a Visiting Scientist and Research Fellow at the Telecommunications and Traffic Control Systems Group of the Delft Technical University, The Netherlands, where he worked on spread spectrum techniques for personal communications systems. From October 1992 to August 1997, he was Senior Research Associate of the Belgian NSF at UCL, and invited Assistant Professor. He is currently a Professor and head of the Institute for Information and Communication Technologies, Electronics and Applied Mathematics. His current interest is in digital communication systems and more precisely resource allocation for OFDM(A)-based multicell systems, MIMO and distributed MIMO, sensor networks, turbo-based communications systems, physical layer security and UWB based positioning.

Dr. Vandendorpe was corecipient of the 1990 Biennal Alcatel–Bell Award from the Belgian NSF for a contribution in the field of image coding. In 2000, he was corecipient (with J. Louveaux and F. Deryck) of the Biennal Siemens Award from the Belgian NSF for a contribution about filter-bank-based multicarrier transmission. In 2004, he was co-winner (with J. Czyz) of the Face Authentication Competition, FAC 2004. He is or has been TPC member for numerous IEEE conferences (VTC Fall, Globecom Communications Theory Symposium, SPAWC, ICC) and for the Turbo Symposium. He was Co-Technical Chair (with P. Duhamel) for the IEEE ICASSP 2006. He was an Editor for Synchronization and Equalization of the IEEE TRANSACTIONS ON COMMUNICATIONS between 2000 and 2002, Associate Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS between 2003 and 2005, and Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING between 2004 and 2006. He was Chair of the IEEE Benelux joint chapter on Communications and Vehicular Technology between 1999 and 2003. He was an elected member of the Signal Processing for Communications committee between 2000 and 2005, and an elected member of the Sensor Array and Multichannel Signal Processing committee of the Signal Processing Society between 2006 and 2008. Currently, he is an elected member of the Signal Processing for Communications committee. He is the Editor-in-Chief for the EURASIP Journal on Wireless Communications and Networking.