MINIMAL ENERGY DECENTRALIZED ESTIMATION BASED ON SENSOR NOISE VARIANCE STATISTICS

Jwo-Yuh Wu, Qian-Zhi Huang, and Ta-Sung Lee
Department of Communication Engineering, National Chiao Tung University,
1001, Ta-Hsueh Road, Hsinchu, Taiwan
Tel: 886-35-712121, ext. 54549; Fax: 886-35-710116
Emails: jywu@cc.nctu.edu.tw; qianzhi.yux@gmail.com; tslee@mail.nctu.edu.tw.

***Abstract*- This paper studies minimal-energy decentralized estimation in sensor networks under best-linear-unbiased-estimator fusion rule. While most of the existing related works require the knowledge of instantaneous noise variances for energy allocation, the proposed approach instead relies on an associated statistical model. The minimization of total energy is subject to certain performance constraint in terms of mean square error (MSE) averaged over the noise variance distribution. A closed-form formula for the overall MSE metric is derived, based on which the problem can be reformulated in the form of convex optimization and is shown to yield an analytic solution. The proposed method shares several attractive features of the existing designs via instantaneous noise variances; through simulations it is seen to significantly improve the energy efficiency against the uniform allocation scheme.**

Index Terms: **Decentralized estimation; Sensor networks; Energy Minimization; Convex optimization; Quantization.**

I. INTRODUCTION

Decentralized estimation has become an important topic in signal processing research for sensor networks [9], [10]. Subject to severe energy and bandwidth limitations, each sensor in this scenario is allowed to transmit only a quantized version of its raw measurement to the fusion center (FC) to generate a final parameter estimate. While quantized messages with longer bit length provide improved data fidelity, the consumed transmission energy is however proportional to the bit loads [3], [8]. As energy efficiency is a critical concern for sensor network design [9], [10], the minimal-energy decentralized estimation problem, formulated in an optimal bit-loading setup, has been recently considered in [3], [6], [8]. One key feature common to these works is that the energy (or bits) allocated to each sensor must be determined via instantaneous local sensor noise characteristics, e.g., the noise variance, if the fusion rule follows the best-linear-unbiased-estimator (BLUE) principle [1]. To improve the estimation performance against the variation of sensing conditions, repeated update of the noise profile would be needed: this comes inevitably at the cost of more training overhead and thus extra energy consumption. One typical approach to resolving such a drawback is to exploit the partial (or long-term) information of the noise characteristics [8]; the related solutions, however, remain yet to be developed.

This paper attempts to provide a solution to minimal-energy decentralized estimation (under BLUE fusion rule) by exploiting long-term noise variance information. We focus on a commonly used statistical model for noise variance, and the estimation performance is assessed through an MSE based metric averaged with respect to the considered distribution. A closed-form expression of the overall MSE requirement is derived, and it is seen to be highly nonlinear in the sensor bit loads. Through analysis the energy-minimization problem is reformulated in the form of convex optimization and is then analytically solved. The proposed optimal scheme shares several interesting aspects pertaining to those based on the instantaneous noise variance information: sensors with bad channel quality (specified via the path distance to FC) are shut off to conserve energy, and for those active nodes the allocated energy is proportional to the individual channel gain. Simulation results show that the proposed optimal solution yields significant energy saving against the equal-bit allocation policy.

II. PRELIMINARY

Consider a wireless sensor network, in which N spatially deployed sensors cooperate with a FC for estimating an unknown deterministic parameter θ. The local observation at the ith node is

$$x_i := \theta + n_i, \; 1 \le i \le N, \tag{2.1}$$

where n_i is a zero-mean measurement noise with variance σ_i^2. Due to bandwidth and power limitations each sensor quantizes its observation into a b_i-bit message, and then transmits this locally processed data to the FC to generate a final estimate of θ. In this paper the uniform quantization scheme with nearest-rounding [4] is adopted; the quantized message at the ith sensor can thus be modeled as

$$m_i := x_i + q_i, \; 1 \le i \le N, \tag{2.2}$$

where q_i is the quantization error which is uniformly distributed with zero mean and variance $\sigma_{q_i}^2 = R^2/(12 \cdot 4^{b_i})$ [4], where $[-R/2, R/2]$ is the available signal amplitude range common to all sensors. With (2.1) and (2.2), the received data from all sensor outputs can be expressed in a vector form as[a]

a. We assume perfect reception of all the messages m_i' at the FC, and the resultant MSE thus serves as a yardstick performance. When the transmission link is modeled as a binary symmetric channel, by following the procedures as in [8] the incurred MSE can be shown to be at most a constant factor away from the benchmark measure, provided that the bit-error-rate is below certain threshold.

$$\underbrace{[m_1 \cdots m_N]^T}_{:=\mathbf{m}} = \underbrace{[1 \cdots 1]^T}_{:=\mathbf{1}} \theta + \underbrace{[n_1 \cdots n_N]^T}_{:=\mathbf{n}} + \underbrace{[q_1 \cdots q_N]^T}_{:=\mathbf{q}}, \quad (2.3)$$

where $(\cdot)^T$ denotes the transpose. This paper focuses on linear fusion rules for parameter recovery. More specifically, by assuming that the noise components $\{\mathbf{n},\mathbf{q}\}$ in (2.3) are mutually independent with covariance matrices $\mathbf{C_n}$ and $\mathbf{C_q}$, the parameter θ is retrieved via the BLUE [1] estimator via

$$\hat{\theta} := \frac{\mathbf{1}^T \mathbf{C}^{-1} \mathbf{m}}{\mathbf{1}^T \mathbf{C}^{-1} \mathbf{1}}, \text{ where } \mathbf{C} := \mathbf{C_n} + \mathbf{C_q}. \quad (2.4)$$

We further assume that the measurement noise n_i's are i.i.d., and the quantization noise q_i's are independent across all sensors; as such the MSE incurred by $\hat{\theta}$ can be immediately computed as [1]

$$E\left|\hat{\theta} - \theta\right|^2 = \left(\mathbf{1}^T \mathbf{C}^{-1} \mathbf{1}\right)^{-1} = \left(\sum_{i=1}^{N} \frac{1}{\sigma_i^2 + R^2 4^{-b_i}/12}\right)^{-1}. \quad (2.5)$$

A commonly used statistical description for sensor noise variance is [3], [8]:

$$\sigma_i^2 = \delta + \alpha z_i, \ 1 \le i \le N, \quad (2.6)$$

where δ models the network-wide noise variance threshold, α controls the underlying variation from the nominal minimum, and $z_i \sim \chi_1^2$ is a central Chi-Square distributed random variable with degrees-of-freedom equal to one [2, p-24]. In the sequel we will exploit the noise variance model (2.6) for minimal-energy decentralized estimation.

III. MAIN RESULTS

A. Problem Formulation

We assume as in [3] that the consumed energy for transmitting the message m_i at the ith sensor is proportional to the number of bits b_i in m_i, that is,

$$E_i = w_i b_i \text{ for some } w_i, \ 1 \le i \le N; \quad (3.1)$$

the energy density factor w_i is defined as [3]

$$w_i := \rho d_i^{\kappa} \cdot \frac{(2^s - 1)}{s} \cdot \ln\left(\frac{4(1-2^{-s})}{sP_b}\right), \quad (3.2)$$

in which ρ is a constant depending on the noise profile, d_i is the distance between the ith node and the FC, κ is the path loss exponent common to all sensor-to-FC links, s is the number of bits per QAM/PSK symbol, and P_b is the target bit error rate. With (3.1), the specification of the energy allocated to the ith sensor thus amounts to determining the number of quantization bits b_i. For a fixed set of noise variances σ_i^2's, the energy minimization problem subject to an allowable parameter distortion level γ (in terms of MSE) can be formulated as:

Minimize $\sum_{i=1}^{N} w_i b_i$, subject to

$$\left(\sum_{i=1}^{N} \frac{1}{\sigma_i^2 + (R^2/12)4^{-b_i}}\right)^{-1} \le \gamma \text{ and } b_i \ge 0, \ 1 \le i \le N, \quad (3.3)$$

or equivalently,

Minimize $\sum_{i=1}^{N} w_i b_i$, subject to

$$\sum_{i=1}^{N} \frac{1}{\sigma_i^2 + (R^2/12)4^{-b_i}} \ge \gamma^{-1} \text{ and } b_i \ge 0, \ 1 \le i \le N. \quad (3.4)$$

To obtain a universal solution irrespective of instantaneous measurement noise conditions, we will consider the following optimization problem, in which the equivalent MSE performance metric in (3.4) is instead averaged with respect to the noise variance statistic characterized in (2.6):

Minimize $\sum_{i=1}^{N} w_i b_i$, subject to

$$\int_{\mathbf{z}} \sum_{i=1}^{N} \frac{1}{\delta + \alpha z_i + (R^2/12)4^{-b_i}} p(\mathbf{z}) \, d\mathbf{z} \ge \gamma^{-1}, \ b_i \ge 0, \ 1 \le i \le N, \quad (3.5)$$

where $\mathbf{z} := [z_1 \cdots z_N]^T$ with $p(\mathbf{z})$ denoting the associated distribution. In (3.5), the constraint that all b_i are nonnegative integers are relaxed to be $b_i \ge 0$ so as to render the problem tractable; once the optimal (real valued) b_i's are computed, the associated bit loads can be obtained via upper integer rounding, as in [3], [8]. The solution to problem (3.5) is discussed next.

B. Proposed Approach

To solve (3.5), a crucial step is to derive an analytic expression of the average MSE performance measure. For this we first note that, since $z_i \sim \chi_1^2$ is i.i.d. and [2, p-24]

$$p_{\chi_1^2}(z) = \frac{1}{\sqrt{2\pi z}} \exp(-z/2) u(z), \quad (3.6)$$

where $u(\cdot)$ is the unit-step function, we have

$$\int_{\mathbf{z}} \sum_{i=1}^{N} \frac{1}{\delta + \alpha z_i + (R^2/12)4^{-b_i}} p(\mathbf{z}) \, d\mathbf{z}$$
$$= \sum_{i=1}^{N} \int_0^{\infty} \frac{1}{\alpha z_i + \underbrace{[\delta + R^2 4^{-b_i}/12]}_{:=\beta_i}} \cdot \frac{e^{-z_i/2}}{\sqrt{2\pi z_i}} dz_i = \frac{1}{\sqrt{2\pi}} \sum_{i=1}^{N} \int_0^{\infty} \frac{e^{-z_i/2}}{(\alpha z_i + \beta_i)\sqrt{z_i}} \, dz_i \quad (3.7)$$

The following lemma (see [7] for a proof) provides a closed-form expression of the integral involved in the summation in (3.7).

Lemma 3.1: With $\alpha > 0$ and $\beta_i > 0$ as defined in (3.7), we have

$$\int_0^{\infty} \frac{e^{-z_i/2}}{(\alpha z_i + \beta_i)\sqrt{z_i}} \, dz_i = \frac{2\pi \cdot e^{\beta_i/2\alpha} \cdot Q\left(\sqrt{\beta_i/\alpha}\right)}{\sqrt{\alpha \beta_i}}, \quad (3.8)$$

where $Q(x) := \int_x^{\infty} \frac{e^{-t^2/2}}{\sqrt{2\pi}} \, dt$ is the Gaussian tail function. □

With (3.7) and (3.8), the optimization problem (3.5) can be equivalently rewritten as

Minimize $\sum_{i=1}^{N} w_i b_i$, subject to

$$\sqrt{\frac{2\pi}{\alpha}} \sum_{i=1}^{N} \frac{e^{\beta_i/2\alpha} \cdot Q\left(\sqrt{\beta_i/\alpha}\right)}{\sqrt{\beta_i}} \ge \gamma^{-1}, \ b_i \ge 0, \ 1 \le i \le N. \quad (3.9)$$

Exact solutions to problem (3.9) appear intractable since the design constraint, in particular, the one accounting for the target MSE, is highly nonlinear in b_i. We will thus seek for suboptimal alternatives which can otherwise admit simple analytic expressions. The underlying approach toward this end is to derive an easy-to-tackle lower bound on the target MSE metric, and then replace the MSE constraint in (3.9) by one which forces the lower bound to be above γ^{-1}: such a procedure will considerably simplify the analysis without incurring any loss in the desired MSE performance. This is done with the aid of the next lemma (see [7] for a proof).

Lemma 3.2: The following inequality holds:

$$\sqrt{\frac{2\pi}{\alpha}}\sum_{i=1}^{N}\frac{e^{\beta_i/2\alpha}\cdot Q\left(\sqrt{\beta_i/\alpha}\right)}{\sqrt{\beta_i}} \geq cNQ\left(\frac{1}{N}\sum_{i=1}^{N}\left(\sqrt{\delta/\alpha}+R2^{-b_i}/\sqrt{12\alpha}\right)\right),$$

where $c := \left(\sqrt{2\pi}e^{\delta/2\alpha}\right)\left(\sqrt{\alpha\left(\delta+R^2/12\right)}\right)^{-1}$. □

The above inequality suggests that it suffices to consider the following modified constraint without incurring any loss in the target MSE:

$$cNQ\left(\frac{1}{N}\sum_{i=1}^{N}\left(\sqrt{\delta/\alpha}+R2^{-b_i}/\sqrt{12\alpha}\right)\right) \geq \gamma^{-1}, \qquad (3.10)$$

or equivalently,

$$\frac{1}{N}\sum_{i=1}^{N}\left(\sqrt{\delta/\alpha}+R2^{-b_i}/\sqrt{12\alpha}\right) \leq Q^{-1}\left(\frac{1}{cN\gamma}\right) \qquad (3.11)$$

since $Q(\cdot)$ is one-to-one and monotone decreasing. Based on the above discussions, we will instead focus on the optimization problem with a modified MSE performance constraint:

Minimize $\sum_{i=1}^{N} w_i b_i$, subject to

$$\frac{R}{\sqrt{12\alpha}N}\sum_{i=1}^{N}2^{-b_i} \leq Q^{-1}\left(\frac{1}{cN\gamma}\right)-\sqrt{\frac{\delta}{\alpha}},\ b_i \geq 0,\ 1 \leq i \leq N. \ (3.12)$$

The main advantage of the proposed formulation is that, in (3.12), the cost function is linear and the constraints are convex; it is thus a convex optimization problem and will moreover lead to a simple closed-form solution as shown below.

C. Optimal Solution

To solve problem (3.12), let us form the Lagrangian as

$$L\left(b_1,\cdots,b_N,\lambda,\mu_1,\cdots\mu_N\right) = \sum_{i=1}^{N} w_i b_i + \lambda\left(\frac{R}{\sqrt{12\alpha}N}\sum_{i=1}^{N}2^{-b_i} - Q^{-1}\left(\frac{1}{cN\gamma}\right)+\sqrt{\frac{\delta}{\alpha}}\right)-\sum_{i=1}^{N}\mu_i b_i \qquad (3.13)$$

the associated set of KKT conditions then reads

$$w_i + \lambda\cdot\frac{(-\ln 2)R2^{-b_i}}{\sqrt{12\alpha}N}-\mu_i = 0,\ 1 \leq i \leq N, \qquad (3.14)$$

$$\lambda\left(\frac{R}{\sqrt{12\alpha}N}\sum_{i=1}^{N}2^{-b_i} - Q^{-1}\left(\frac{1}{cN\gamma}\right)+\sqrt{\frac{\delta}{\alpha}}\right) = 0, \qquad (3.15)$$

$$\lambda \geq 0,\ \mu_i \geq 0,\ \mu_i b_i = 0,\ b_i \geq 0,\ 1 \leq i \leq N. \qquad (3.16)$$

The optimal solution can be obtained by solving (3.14)~(3.16), and is given by the following theorem (see [7] for detailed derivation).

Theorem 3.3: Assume $w_1 \geq w_2 \geq \cdots \geq w_N$ without loss of generality, and define the function

$$f(K) := \left(Nw_K\right)^{-1}\sum_{i=1}^{K} w_i,\ 1 \leq K \leq N. \qquad (3.17)$$

Let $1 \leq K_1 \leq N$ be such that $f(K_1-1) < 1$ and $f(K_1) \geq 1$. Then we have

$$b_i^{opt} = \begin{cases} 0, & 1 \leq i \leq K_1-1, \\ \log_2\left\{\dfrac{R\lambda^{opt}}{\sqrt{12\alpha}Nw_i}\right\}, & K_1 \leq i \leq N, \end{cases} \qquad (3.18)$$

where $\lambda^{opt} = \left[Q^{-1}\left(c^{-1}N^{-1}\gamma^{-1}\right)-\sqrt{\delta/\alpha}\right]^{-1}\sum_{j=1}^{K_1} w_j$. □

D. Discussions

1. Since $0 \leq b_i < \infty$, a necessary condition for validating the MSE constraint in (3.12) is therefore
$$Q^{-1}\left(\frac{1}{cN\gamma}\right)-\sqrt{\frac{\delta}{\alpha}} \geq 0, \text{ or } \frac{1}{cN\gamma} \leq Q\left(\sqrt{\frac{\delta}{\alpha}}\right) \qquad (3.19)$$
because $Q(\cdot)$ is one-to-one and monotone decreasing. By definition of the constant c and with (3.19), the MSE attainable by the proposed method is lower bounded by
$$\gamma \geq \left[Ne^{\delta/2\alpha}Q\left(\sqrt{\delta/\alpha}\right)\sqrt{\frac{2\pi}{\alpha(\delta+R^2/12)}}\right]^{-1}. \qquad (3.20)$$
2. Recall from (3.2) that the energy density factor w_i is proportional to the path loss gain d_i^{κ} (if the same bit error rate is assumed throughout all the links). Large values of w_i, in particular, correspond to sensors deployed far away from the FC (with large d_i), usually with poor background channel gains. In light of this point, the proposed optimal solution (3.18) is intuitively attractive: sensors associated with the (K_1-1)th largest w_i's are turned off to conserve energy. We note that a similar energy conservation strategy via shutting off sensors alone poor channel links is also found in [8], in which an energy model with exponential dependency on b_i is otherwise adopted and a scenario with instantaneous noise variances available to the FC is considered.
3. We further note from (3.18) that, for those active nodes, the assigned message length is inversely proportional to w_i: this is intuitively reasonable since sensors with better link conditions should be allocated with more bits (energy) to realize the desired network-wide performance.
4. Based on the inequality constraint for MSE in (3.12), the equal-bit scheme maintaining the desired MSE can be obtained by obtained as
$$\tilde{b} = \log_2\left\{\frac{R}{\sqrt{12\alpha}\left[Q^{-1}\left(c^{-1}N^{-1}\gamma^{-1}\right)-\sqrt{\delta/\alpha}\right]}\right\}. \qquad (3.21)$$
Simulation results in the next section show that the proposed optimal scheme (3.18) yields significant energy saving when compared with (3.21).

IV. NUMERICAL SIMULATION

This section illustrates through numerical simulation the energy saving efficiency of the proposed solution (3.18) over the uniform allocation scheme (3.21). For a fixed set of energy density factors w_i , $1 \le K \le N$, the performance is measured via the percentage of energy saving (PES) as defined in [3], [8]. In each trial, we simply set $w_i = d_i^{\kappa}$, where $\kappa = 3.5$ and $d_i = 10 + 10Z_i$, with $Z_i \sim \chi_1^2(z)$ being i.i.d.; the results are averaged over 50,000 independent trials. The total number of sensors is $N = 1500$, under $\gamma = 0.005$. For minimal noise variance threshold fixed at $\delta = 0.8$, Figure 1-(a) shows the PES for $0.1 \le \alpha \le 1.6$, and Figure 1-(b) depicts the computed $\tilde{b}$ in (3.21). We can observe that the PES exhibits two "jumps": this accounts for the two level changes of $\tilde{b}$ as α varies. Also, within each duration of constant $\tilde{b}$, energy efficiency of the optimal solution (3.18) improves as α increases (a large α corresponds to a more inhomogeneous sensing environment). We note that a similar phenomenon has been observed in the existing works relying on instantaneous noise variance knowledge [3], [8], owing to the fact that, as the sensing condition becomes more inhomogeneous, it is more likely that a large fraction of sensors suffers from poor measurement quality and will thus be shut off, leading to improved energy efficiency. Since the proposed solution (3.18) (based on statistical noise variance description) would reflect the long-term characteristics of the schemes [3], [8], this consistency is thus expected. We repeat the experiment by fixing $\alpha = 0.4$ and varying the minimal threshold δ ; the results are shown in Figure 2. As we can see, the PES exhibits a counter tendency as compared to Figure 1: for each duration of constant $\tilde{b}$ the energy saving achieved by solution (3.18) is nonetheless lowered as δ increases. This is reasonable since large δ 's result in severe noise corruption in *all* sensor measurements: more sensor nodes should thus be turned on (thus potentially more energy consumption) to provide a sufficient amount of information for MSE reduction. Overall we also conclude from the figures that the proposed optimal solution is capable of reducing about 80% energy consumption when compared with the uniform-allocation scheme; the energy saving efficiency is particularly significant when the minimal variance threshold is small or the variation factor is large.

REFERENCES

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

[2] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*, Prentice-Hall PTR, 1998.

[3] A. Krasnopeev, J. J. Xiao, and Z. Q. Luo, "Minimum energy decentralized estimation in sensor network with correlated sensor noise," *Proc. ICASSP05'*, vol. 3, pp. 673-676.

[4] S. J. Orfanidis, *Introduction to Signal Processing*, Prentice-Hall, Inc., 1996.

[5] M. K. Simon and M. S. Alouini, *Digital Communication over Fading Channels: A Unified Approach to Performance Analysis*, John Wiley & Sons, Inc., 2000.

[6] P. Venkitasubramaniam, G. Mergen, L. Tong, and A. Swami, "Quantization for distributed estimation in large scale sensor networks," *Proc. ICISIP 2005*, pp. 121-127.

[7] J. Y. Wu, Q. Z. Huang, and T. S. Lee, "Minimal energy decentralized estimation based on sensor noise variance statistics," *IEEE Trans. Signal Processing*, submitted for publication.

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

[9] J. J. Xiao, A. Ribeiro, Z. Q. Luo, and G. B. Giannakis, "Distributed compression-estimation using wireless sensor networks," *IEEE Signal Processing Magazine*, vol. 23, no. 4, pp. 27-41, July 2006.

[10] Q. Zhao, A. Swami, and L. Tong, "The interplay between signal processing and networking in sensor networks," *IEEE Signal Processing Magazine*, vol. 23, no. 4, pp. 84-93, July 2006.

Acknowledgment: This work is sponsored jointly by the National Science Council under grant NSC-96-2752-E-002-009, by the Ministry of Education of Taiwan under the MoE ATU Program, and by MediaTek research center at National Chiao Tung University, Taiwan.

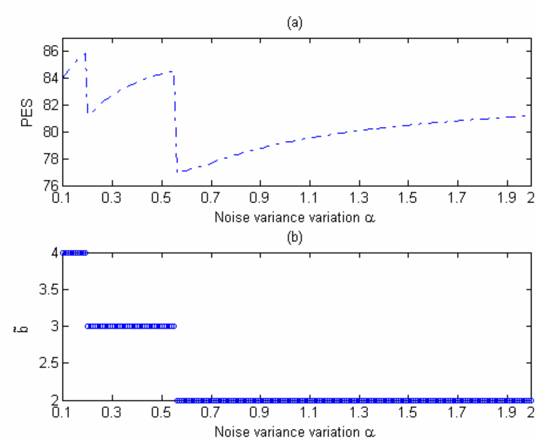

Figure 1. (a) Percentage of energy saving; (b) Bit number of equal-energy scheme ($\delta = 0.8$).

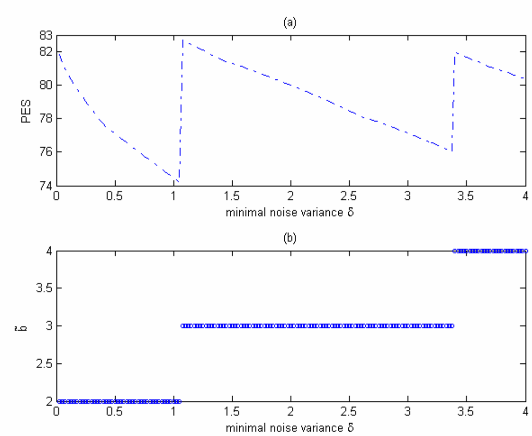

Figure 2. (a) Percentage of energy saving; (b) Bit number of equal-energy scheme ($\alpha = 0.4$).