ROBUST POWER ALLOCATION FOR DISTRIBUTED ESTIMATION IN WIRELESS SENSOR NETWORKS

Umar Rashid, Fahim Gohar Awan

Department of Electrical, Electronics and Telecommunication Engineering, UET, KSK Campus, Lahore, Pakistan

Contact: fahimgohar@hotmail.com

ABSTRACT *— Problem considered in this paper is distributed estimation of an unknown parameter using wireless sensor network (WSN) in a random sensing environment. The system model consists of distributed sensor nodes and a central fusion center (FC). The estimation of the unknown parameter is carried out at the FC using a minimum mean square error (MMSE) estimator. It is assumed that the exact knowledge of the sensing model (local channel) is not available at the FC. A robust power allocation algorithm is then proposed to accommodate the uncertainties in local channel knowledge. Furthermore, performance evaluation of the distributed estimation is investigated with respect to network scaling (i.e., network size) by statistically averaging out the effects of the random sensing gain and then finally expressing the mean-squared error (MSE) of the estimate as a function of the network size.*

Index Terms —Wireless sensors, Fusion center, Concave-convex procedure.

I. INTRODUCTION

Network based on wireless sensors have observations that are correlated spatially and partially of the underlying source. The correlation exists for a bounded geographic region and placement of sensors within its boundaries eg., same event recording with acoustic sensors. On the other hand, communication channel and noise measurements may not observe similar channel conditions across these sensors. Thus uniform power allocation strategy in communication is not optimal. In this paper, we investigate the problem of optimal network power allocation adaptively within its constraints in order to optimize the MSE reconstruction. The adaptive power allocations at the FC are jointly estimated and through feedback channels transmitted to the relevant sensors.

Under the assumptions of unknown local sensors' channels, we propose a method to find the number of sensors required to achieve a given target distortion D with fixed total sensor power budget P_{tot}. Here, we consider the distributed estimation problem for sensor networks from a scaling law perspective.

Due to the difficulty of having the exact knowledge of local sensors channel information, the power allocation scheme should also take into account the channel estimation errors. Among recent works, power allocation is considered with uncertainty in noise variance in a wireless sensor network [1]. Robust power allocation for location-aware networks has been addressed in [2]. Unlike the work in [3], where an SDP is established to solve optimal power allocation problem with deterministically known channel gains, the current draft considers the optimal power allocation as maximization of P_r $(MSE \leq \epsilon)$ where the randomness in MSE arises out of random channel gains. So the issue is to derive a deterministic expression for the aforementioned probability and then optimize it over the unknown $\{\alpha_i\}_{i=1}^N$subject to total transmit power constraint.

Notations: In this paper, italic bold-faced letters wth upper and lower case denotes matrices and random vectors, respectively, while upright letters stand for their realizations. *f*() and $\mathbb{E}\{.\}$ represent probability density function and expectation, respectively. b_i represent the vector $\boldsymbol{b}$ i^{th} element, $B(i, j)$ is the (i, j)th element in $\boldsymbol{B}$ matrix, its determinant $|\boldsymbol{B}|$, i^{th} column of $N \times N$ identity matrix I_N is e_i. $(\,)^H$ denotes complex-conjugate transposition. Base e is used for logarithmic computing. $x \sim CN(\boldsymbol{m}, \boldsymbol{R})$ represents the complex Gaussian distributed random variable *x with* circular symmetry having variance *R* and m mean.

II. SYSTEM MODEL

Let a sensor network consisting of N nodes make independent measurements of a zero-mean random scalar parameter $\theta \sim (0, \sigma_\theta^2)$. Under a noisy environment, observation equation at the i^{th} takes the following form:

$$x_i = h_i\,\theta + n_i \tag{1}$$

where h_i is the i^{th} local channel gain, and $n_i \sim (0, \sigma_{i,n})$ is the local AWGN. These measurements are independently transmitted over orthogonal channels by the sensors after multiplying with an amplification factor $\{\alpha_i\}_{i=1}^N$ to a fusion center (FC) which admits the following equation:

$$\boldsymbol{z} = \boldsymbol{D}_h\,\boldsymbol{g}\theta + \boldsymbol{w} \tag{2}$$

where $\boldsymbol{D}_h = \text{diag}[h_1\alpha_1, \dots, h_N\alpha_N]$, $\boldsymbol{g} = [g_1, g_2, \dots, g_N]^{\boldsymbol{T}}$ is the channel gain from sensor network to the FC, and $\boldsymbol{w} = [g_1\alpha_1 n_1 + v_1, \dots, g_N\alpha_N n_N + v_N]^{\boldsymbol{T}}$ normally distributed as $\boldsymbol{w} \sim N(0, \boldsymbol{R}_w)$ with $\boldsymbol{R}_w = \text{diag}[g_i^2\alpha_i^2\sigma_{i,n}^2 + \zeta_i^2]$. Consider the case when h_i is modeled as a Gaussian random variable, i.e., $h_i \sim N(0, \sigma_{h,i}), i = 1, \dots, N.$ Upon receiving these observations, the FC carries out LMMSE estimate given by:

$$\hat{\theta} = \bar{\theta} + \boldsymbol{R}_{\theta z}\boldsymbol{R}_z^{-1}(\boldsymbol{z} - \bar{\boldsymbol{z}})$$

and the expression for MSE takes the form

$$\begin{aligned}\mathbb{E}[(\theta-\hat{\theta})^2] &= \sigma_\theta^2 - \boldsymbol{R}_{\theta z}\boldsymbol{R}_z^{-1}\boldsymbol{R}_{\theta z}^T \\ &= \sigma_\theta^2 - \sigma_\theta^4\boldsymbol{h}^T\boldsymbol{\mathfrak{D}}_g^T(\boldsymbol{\mathfrak{D}}_g\boldsymbol{h}\sigma_\theta^2\boldsymbol{h}^T\boldsymbol{\mathfrak{D}}_g^T + \boldsymbol{R}_w)^{-1}\boldsymbol{\mathfrak{D}}_g\boldsymbol{h} \\ &= \left(\sigma_\theta^{-2} + \boldsymbol{h}^T\boldsymbol{\mathfrak{D}}_g^T\boldsymbol{R}_w^{-1}\boldsymbol{\mathfrak{D}}_g\boldsymbol{h}\right)^{-1}\end{aligned} \tag{3}$$

where $\boldsymbol{\mathfrak{D}}_g = \text{diag}\left[\alpha_i g_i\right]$

If $\boldsymbol{B} = \text{diag}\left[\frac{g_1^2\alpha_1^2}{\alpha_1^2 g_1^2\sigma_{1,n}^2+\zeta_1^2}, \dots, \frac{g_N^2\alpha_N^2}{\alpha_N^2 g_N^2\sigma_{N,n}^2+\zeta_N^2}\right] = \text{diag}\left[\beta_1, \beta_2, \dots, \beta_N\right]$, then

$$M = \mathbb{E}[(\theta-\hat{\theta})^2] = \frac{1}{\sigma_\theta^{-2} + ||\boldsymbol{h}||_{\boldsymbol{B}}^2} \tag{4}$$

Because of the random nature of the MSE expression, we present its CDF under the assumption of $\boldsymbol{h} \sim CN(0, I)$:

$$F_M(m) = \sum_{i=1}^N \frac{\beta_i^{N-1}}{\prod_{l\neq i}^N(\beta_i - \beta_l)} e^{-\frac{(1-\sigma_\theta^{-2}m)}{\beta_i m}}\left[\tilde{u}(m) - \tilde{u}(1-\sigma_\theta^{-2}m)\right] \tag{5}$$

where $\tilde{u}(.)$ is the step function. First moment of M is:

$$\mathbb{E}[M] = \frac{1}{\sigma_\theta^{-2}} - \frac{1}{\sigma_\theta^{-2}} \sum_{i=1}^{N} \frac{\beta_i^{N-1} e^{\frac{\sigma_\theta^{-2}}{\beta_i}}}{\prod_{l\neq i}^{N}(\beta_i - \beta_l)} \mathrm{E}_2\left(\frac{\sigma_\theta^{-2}}{\beta_i}\right) \quad (6)$$

Note that in most of the cases, exact channel gains $\{h_i\}_{i=1}^{N}$ are not known at the FC, and instead it has to rely on the prior probability distribution of the channel, or find some estimate of the channel. In the latter case, the actual channel can be modeled as a sum of the estimation error $\Delta\bar{h}_i$ and the estimate $\hat{h}_i$,

$$h_i = \hat{h}_i + \Delta\bar{h}_i \quad (7)$$

where $\Delta\bar{h}_i \sim (0, \bar{\delta}_i^2)$ is the estimation error for the i^{th} sensor's local channel. One possible methodology for power-scheduling is by replacing h_i by its estimate $\hat{h}_i$ in the formulations of the resulting MSE. This ignores the channel estimate error therefore it is a naive-approach. In the former case, we resort to the following approach to design amplification factors.

III. OPTIMUM POWER ALLOCATION

In this method, we maximize $\mathbb{E}\{\|\boldsymbol{h}\|_{\boldsymbol{B}}^2\} = \sum_{i=1}^{N} |h_i|^2 \boldsymbol{B}(i,i)$. It should be remembered that $\boldsymbol{h}$ is assumed to be characterized by its first and second order moments, and its exact value is not available at the FC. Assume that the channel is normally distributed with zero-mean and unity variance. Under these assumptions, we formulate the following power allocation optimization problem

$$\max_{\boldsymbol{\gamma}\in\mathcal{R}_+^N} \sum_{i=1}^{N} \frac{g_i^2\gamma_i}{g_i^2\gamma_i\sigma_{i,n}^2 + \zeta_i^2} \quad (8a)$$

$$\text{s.t.} \quad \langle \Sigma_\gamma \boldsymbol{R}_x \rangle \le P_{total} \quad (8b)$$

Where $\alpha_i^2 = \gamma_i$, $\Sigma_\gamma = \mathrm{diag}\{\gamma_i\}$, and $\boldsymbol{R}_x = \sigma_\theta^2\mathbb{E}\{\boldsymbol{h}\boldsymbol{h}^T\} + \boldsymbol{R}_n$. The objective function in (8) is a sum of linear fractional functions and is a nonconvex function [4] for which, in general, there is one or more locally optimal solutions. It is, therefore, difficult to find globally optimal solution for such problems. However, through efficient implementation of concave-convex procedure [5], also known as d.c. (two convex/concave sets difference) programming [6], an iterative algorithm can be developed of very low complexity to find the solution.

First we express the maximization program (8) in the canonical form as [6]:

$$\max_{\boldsymbol{z}}[f(\boldsymbol{z}) - g(\boldsymbol{z})] \ : \ \boldsymbol{z} \in \mathcal{K} \quad (9)$$

where K is a convex and compact set, $\boldsymbol{z}$ denotes a vector variable and/or matrix, $f(.)$ and $g(.)$ are a concave and smooth function. Suppose that at $z^{(k)}$ $z^{(\kappa)} \in \mathcal{K}$ and $\nabla g(z^{(\kappa)})$ is the gradient of $g(.)$. Then [6]

$$f(\boldsymbol{z}) - g(\boldsymbol{z}) \ge f(\boldsymbol{z}) - g(\boldsymbol{z}^{(\kappa)}) - \langle \nabla g(\boldsymbol{z}^{(\kappa)}), \boldsymbol{z} - \boldsymbol{z}^{(\kappa)} \rangle \quad \forall \boldsymbol{z} \in \mathcal{K}$$

then for d.c. program (9) global lower bound maximization are provided by the concave program as follows:

$$\max_{\boldsymbol{z}}[f(\boldsymbol{z}) - g(\boldsymbol{z}^{(\kappa)}) - \langle \nabla g(\boldsymbol{z}^{(\kappa)}), \boldsymbol{z} - \boldsymbol{z}^{(\kappa)} \rangle] \ : \ \boldsymbol{z} \in \mathcal{K} \quad (10)$$

where the feasible solution is $z^{(\kappa)}$. whereas for $z^{(\kappa+1)}$ optimal solution of (10) $= f(z^k) - g(z^k)$,

$$\begin{aligned} f(\boldsymbol{z}^{(\kappa+1)}) - g(\boldsymbol{z}^{(\kappa+1)}) &\ge f(\boldsymbol{z}^{(\kappa+1)}) - g(\boldsymbol{z}^{(\kappa)}) - \langle \nabla g(\boldsymbol{z}^{(\kappa)}), \boldsymbol{z}^{(\kappa+1)} - \boldsymbol{z}^{(\kappa)} \rangle \\ &\ge f(\boldsymbol{z}^{(\kappa)}) - g(\boldsymbol{z}^{(\kappa)}) - \langle \nabla g(\boldsymbol{z}^{(\kappa)}), \boldsymbol{z}^{(\kappa)} - \boldsymbol{z}^{(\kappa)} \rangle \\ &= f(\boldsymbol{z}^{(\kappa)}) - g(\boldsymbol{z}^{(\kappa)}), \end{aligned}$$

thus (10) converges to an optimal solution using path-following algorithm.

Let $\mathcal{F}(\gamma) = \sum_{i=1}^{N} \frac{g_i^2\gamma_i}{\gamma_i^2\sigma_{i,n}^2 + \zeta_i^2}$. Since for any practical scenario $F(\gamma)$ is always nonnegative, it follows that it can be replaced with $G(\gamma) = \log(F(\gamma))$, where $\log(x)$ is a monotonically increasing and a concave function of $x > 0$. Furthermore, by Jensen's Inequality, for any concave function $\psi(X)$, we have $\psi(\mathbb{E}\{X\}) \ge \mathbb{E}\psi(X)\}$. Hence, we express the given objective function in terms of its lower bound which is then maximized to obtain sub-optimal solutions for the amplification factors. So we have

$$\mathcal{G}(\gamma) \ge \sum_{i=1}^{N} \varphi_i(\gamma_i)$$

where

$$\begin{aligned} \varphi_i(\gamma_i) &= \log(g_i^2\gamma_i) - \log(g_i^2\gamma_i + \zeta_i^2) \\ &= f_{i,01}(\gamma_i) - f_{i,02}(\gamma_i) \end{aligned}$$

where both $f_{i,01}(\gamma_i)$ and $f_{i,02}(\gamma_i)$ are concave functions of γ_i. Thus, $\varphi_i(\gamma_i)$ is a difference of two concave functions. Since the sum of d.c. functions preserves its d.c. structure, (8) is equivalent to the following d.c. program

$$\max_{\boldsymbol{\gamma}\in\mathcal{R}_+^N} \sum_{i=1}^{N} \varphi_i(\gamma_i) \quad : \quad (8b)$$

$(\gamma^{(K+1)})$ the path following initialized from a feasible solution (γ^0) of (8), represents the optimal solution as:

$$\max_{\boldsymbol{\gamma}\in\mathcal{R}_+^N} \sum_{i=1}^{N} \left[f_{i,01}(\gamma_i) - f_{i,02}(\gamma_i^{(\kappa)}) - \langle \nabla f_{i,02}(\gamma_i^{(\kappa)}), \gamma_i - \gamma_i^{(\kappa)} \rangle \right] : \quad (8b) \qquad (11)$$

where

$$\langle \nabla f_{i,02}(\gamma_i^{(\kappa)}), \gamma_i - \gamma_i^{(\kappa)} \rangle = \frac{\gamma_i - \gamma_i^{(\kappa)}}{g_i^2\gamma^{(\kappa)} + \zeta_i^2}$$

Algorithm 1 sketches the implementation of DCI canonical d.c program.

AlgorithmI

initialization: Choose an initial $\boldsymbol{\mathcal{X}}^{(0)} \in \mathcal{D}$

calculate $\gamma^{(0)} := F_{01}(\boldsymbol{\mathcal{X}}^{(0)}) - F_{02}(\boldsymbol{\mathcal{X}}^{(0)})$

Set *k = 0*

Repeat

Solve the convex program (11)

to obtain the optimal solution $\boldsymbol{\mathcal{X}}^{(\kappa+1)}$

Calculate $\gamma^{(\kappa+1)} := F_{01}(\boldsymbol{\mathcal{X}}^{(\kappa+1)}) - F_{02}(\boldsymbol{\mathcal{X}}^{(\kappa+1)})$

Set $\kappa = \kappa + 1$ **until** $\frac{\gamma^{(\kappa+1)} - \gamma^{(\kappa)}}{\gamma^{(\kappa)}} \le \epsilon$

for a prescribed tolerance ϵ

IV. ASYMPTOTIC ANALYSIS AND SCALING OF SENSOR NETWORK

In the previous section, we focused on maximizing $\mathbb{E}\{\|\boldsymbol{h}\|_{\boldsymbol{C}}\}$ where local channel gains were assumed to be normally distributed random variables. On the other hand, channel gain vector $\boldsymbol{g}$ is equally likely to be unavailable to the FC, and instead it is only characterized by its statistical distribution. In order to circumvent the randomness of $\boldsymbol{g}$, we average out the effects of randomly varying channel gain for a very large number of sensor nodes, i.e., when $N\rightarrow\infty$. We have

$$\mathbb{E}\{\|\boldsymbol{h}\|_{\boldsymbol{B}}^2\} := \mathcal{S} = \sum_{n=1}^{N} \frac{\sigma_h^2 g_i^2 \alpha_i^2}{g_i^2 \alpha_i^2 \sigma_{i,n}^2 + \zeta_i^2}$$

Assuming equal power allocation among sensor nodes, i.e., $\alpha_i = \sqrt{\frac{P_T}{N(\sigma_\theta^2 + \sigma_{i,n}^2)}}$, we have

$$\mathcal{S} = \sum_{n=1}^{N} \frac{\vartheta_i P_T}{\vartheta_i P_T \sigma_{i,n}^2 + N\sigma_\theta^2 \zeta_i^2} \quad (12)$$

where $\vartheta_i = \sigma_h^2 g_i^2$. With the growing numbers of sensor nodes, we desire to observe the distortion behavior asymptotically without bounds. We start with

$$\frac{\vartheta_i P_T}{N\sigma_\theta^2\zeta_i^2} - \frac{\vartheta_i P_T}{\vartheta_i P_T \sigma_{i,n}^2 + N\sigma_\theta^2 \sigma_{i,h}^2 \zeta_i^2} = \frac{\vartheta_i^2 P_T^2 \sigma_{i,n}^2}{N\sigma_\theta^2 \zeta_i^2(\vartheta_i P_T \sigma_{i,n}^2 + N\sigma_\theta^2\zeta_i^2)} \leq \frac{\vartheta_i^2 P_T^2 \sigma_{i,n}^2}{N^2}$$

Thus, we have the following inequalities:

$$\frac{\vartheta_i P_T}{N\sigma_\theta^2\zeta_i^2} - \frac{\vartheta_i^2 P_T^2 \sigma_{i,n}^2}{N^2} \leq \frac{\vartheta_i P_T}{\vartheta_i P_T \sigma_{i,n}^2 + N\sigma_\theta^2 \sigma_{i,h}^2 \zeta_i^2} \leq \frac{\vartheta_i P_T}{N\sigma_\theta^2\zeta_i^2}$$

According to (12), we have

$$\sum_{n=1}^{N} \frac{\vartheta_i P_T}{N\sigma_\theta^2\zeta_i^2} - \sum_{n=1}^{N} \frac{\vartheta_i^2 P_T^2 \sigma_{i,n}^2}{N^2} \leq \mathcal{S} \leq \sum_{n=1}^{N} \frac{\vartheta_i P_T}{N\sigma_\theta^2\zeta_i^2}$$

According to the strong Law of Large Numbers for the case when $N\rightarrow\infty$, we have

$$\sum_{n=1}^{N} \frac{\vartheta_i P_T}{N\sigma_\theta^2\zeta_i^2} \rightarrow \frac{P_T}{\sigma_\theta^2}\mathbb{E}\{\vartheta_1/\zeta_1^2\}, \quad \sum_{n=1}^{N} \frac{\vartheta_i^2 P_T^2 \sigma_{i,n}^2}{N^2} \rightarrow 0$$

Therefore

$$\lim_{N\rightarrow\infty} \mathcal{S} = \frac{P_T}{\sigma_\theta^2}\mathbb{E}\{\vartheta_1/\zeta_1^2\} \quad (13)$$

which implies that according to (4), under the limiting case of infinite number of sensor nodes for LMMSE estimator, MSE is given by

$$\lim_{N\rightarrow\infty} \mathbb{E}\{(\theta - \hat{\theta})^2\} = \frac{1}{\sigma_\theta^{-2} + \frac{P_T}{\sigma_\theta^2}(\sigma_g \sigma_h/\zeta_1)^2} \quad (14)$$

V. SIMULATION RESULTS

With the help numerical simulations, this section confirms the analytic results obtained in the previous sections. These numerical results range from the high SNR channel conditions for sensor network performance down to the relatively poorer channel conditions. By drawing comparison between the simulations and the analytical results developed in previous section, we summarize the robust power allocation along with network scaling characteristics and explore the system parameters that affect the performance when the network scales.

For a sensor network that consists of N = 20 slow-power, wirelessly networked sensors scattered randomly across a region. It is required to estimate a random complex signal of unit variance under Rayleigh channel fading conditions. These channel conditions are similarly assumed also for the orthogonal channels between the sensors and the FC. $\sigma_h^2 = 1$ is the channel fading variance. Both local and global noise variances are set as $\sigma_n^2 = \zeta^2 = 0.01$ while the source is assumed to have unit variance $\sigma_\theta^2 = 1$. The simulation curves in the following figures (1-4) are simulated and analytically compared by taking the means of results of 1000 independent runs.

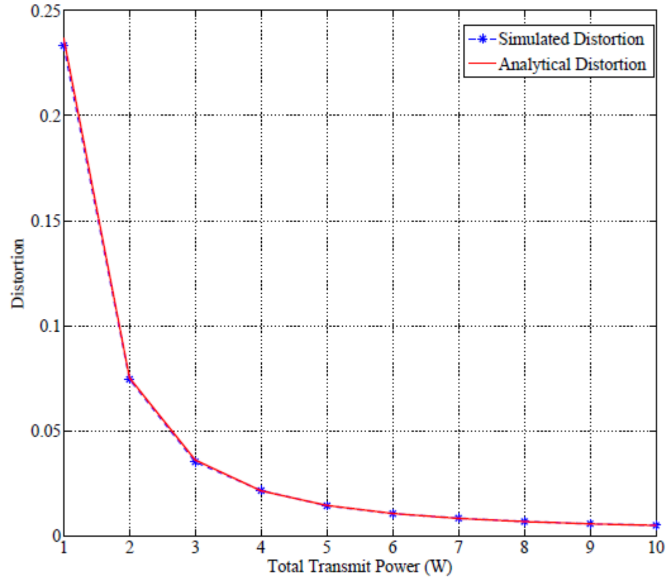

Figure 1: Estimation distortion plotted as a function of total transmit power budget through empirical (6) and analytical (4) evaluation of mean-squared error for N = 10

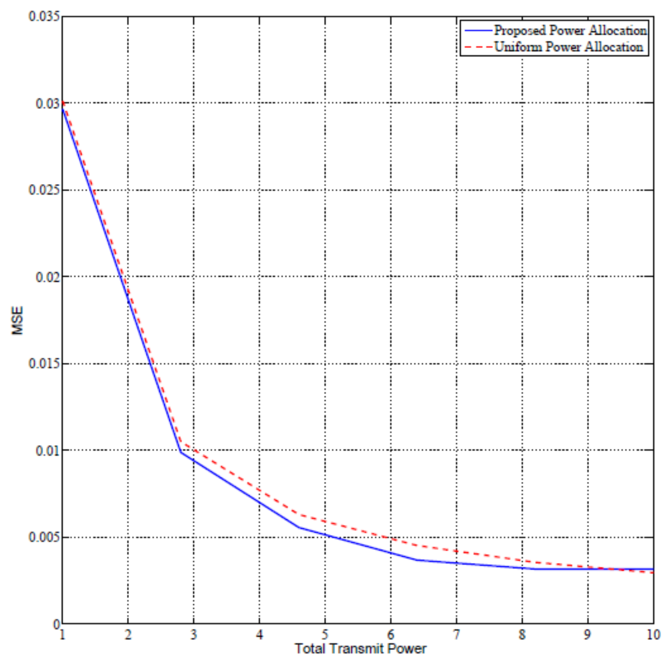

Figure 2: MSE estimation performance comparison between uniform power allocation and the allocation strategy proposed in Algorithm 1 for N = 20

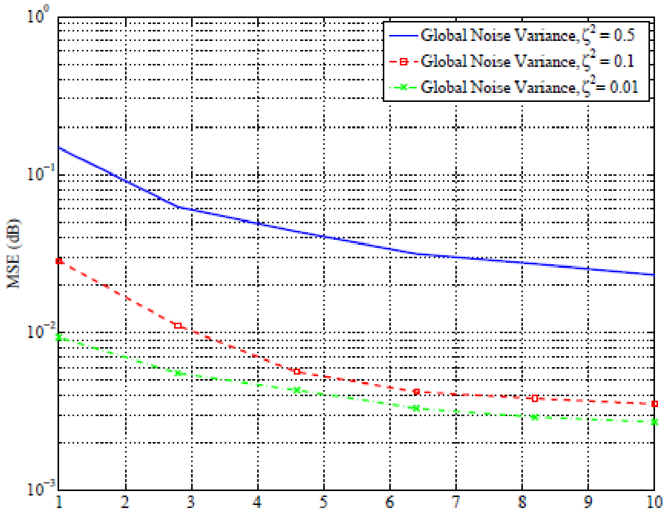

Figure 3: MSE estimation performance vs. total transmit power (W) for various values of noise variances when (11) is implemented for N = 10

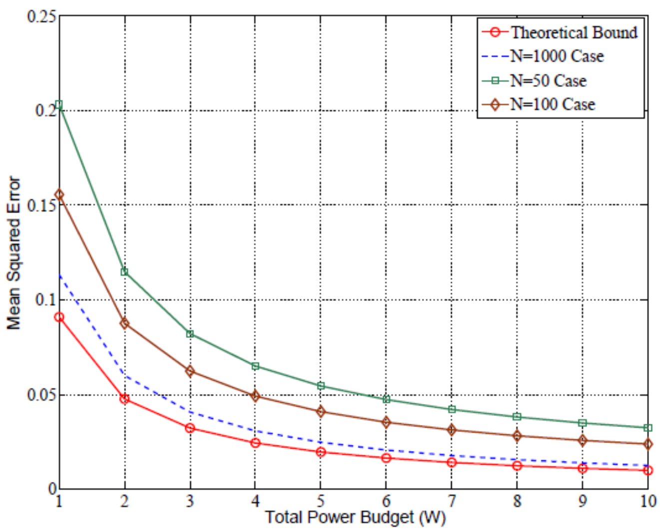

Figure 4: Asymptotic mean-squared error comparison for sufficiently large-sized sensor network with the theoretical bound derived in (14).

VI. CONCLUSION

This paper addresses the power allocation problem for a sensor network when there is a random local sensor channel with no exact value. It makes the task of estimation difficult at the FC. We propose an efficient strategy based upon d.c. programming to circumvent the randomness of the local channel uncertainty. Furthermore, under the same random environment we present the scaling behavior of the network as its size becomes very large.

REFERENCES

[1] J. Y. Wu and T. Y. Wang, “Power allocation for robust distributed best-linear-unbiased estimation against sensing noise variance uncertainty,” *IEEE Transactions on Wireless Communications*, vol. 12, pp. 2853–2869, June 2013.

[2] W. W. L. Li, Y. Shen, Y. J. Zhang, and M. Z. Win, “Robust power allocation for energy-efficient location-aware networks,” *IEEE/ACM Transactions on Networking*, vol. 21, pp. 1918–1930, Dec. 2013.

[3] U. Rashid, H. D. Tuan, H. H. Kha, and P. Apkarian, “Globally optimized power allocation in multiple sensor fusion for linear and nonlinear networks,” *IEEE Transactions on Signal Processing*, vol. 60, no. 2, pp. 903–915, 2012.

[4] S. Schaible, “A note on the sum of a linear and linear-fractional function,” *Naval Research Logistics Quarterly*, vol. 24, no. 4, pp. 691–693, 1977.

[5] A. L. Yuille and A. Rangarajan, “The concave-convex procedure,” *Neural Computation*, vol. 15, no. 4, pp. 915–936, 2003.

[6] H. Tuy, *Convex Analysis and Global Optimization*. Kluwer Academic, 1998.