

On the Effect of Correlated Measurements on the Performance of Distributed Estimation

Mohammed F. A. Ahmed[†], Tareq Y. Al-Naffouri^{†,‡}, and Mohamed-Slim Alouini[†]

[†] King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia.

[‡] King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia.

Email: {m.ahmed, tareq.alnaffouri, slim.alouini}@kaust.edu.sa

Abstract—We address the distributed estimation of an unknown scalar parameter in Wireless Sensor Networks (WSNs). Sensor nodes transmit their noisy observations over multiple access channel to a Fusion Center (FC) that reconstructs the source parameter. The received signal is corrupted by noise and channel fading, so that the FC objective is to minimize the Mean-Square Error (MSE) of the estimate. In this paper, we assume sensor node observations to be correlated with the source signal and correlated with each other as well. The correlation coefficient between two observations is exponentially decaying with the distance separation. The effect of the distance-based correlation on the estimation quality is demonstrated and compared with the case of unity correlated observations. Moreover, a closed-form expression for the outage probability is derived and its dependency on the correlation coefficients is investigated. Numerical simulations are provided to verify our analytic results.

I. INTRODUCTION

Recent advances of micro-sensor fabrication technology allow for producing cheap and small sensor nodes with wireless communication capabilities. Consequently, Wireless Sensor Networks (WSNs) become an economically sound solution to wide range of applications such as environmental and wildlife habitat monitoring, target tracking for defense purposes, and health care [1]. Typical WSN consists of large number of sensor nodes deployed in an area of interest to collect specific information about the surrounding environment. The need for large number of sensor nodes in WSNs while being cost-effective constraints the industry standards to produce battery-powered sensor nodes with simple hardware. As a result of the limited energy and processing capabilities of sensor nodes, the collected information has to be sent to a Fusion Center (FC) for centralized processing.

One important application of WSNs is the distributed estimation of scalar parameters (see, e.g., [2], and references therein). In such application, sensor nodes transmit their observations over a Multiple Access Channel (MAC) to the FC. The received signal is distorted by the channel fading and the additive noise. The FC is required to reconstruct the source parameter with minimum Mean-Square Error (MSE). Depending on the available information about the source statistics, different estimators can be used to achieve the MSE criterion. The performance of the Best Linear Unbiased Estimation (BLUE) [3], Minimum Mean Squared Error (MMSE) estimator [4], [5], and Maximum Likelihood Estimator (MLE) [6], [7] are studied in literature. Both orthogonal

MAC [8] and coherent MAC [9], [10] are considered in the distributed estimation problem. Assuming Gaussian source signal and noise, amplify-and-forward schemes significantly outperform the traditional source-channel coding for both multiple access channels [11]. Optimal power allocation for sensor nodes under different constraints is addressed in [3], [4], [8]. Asymptotic behavior of the distortion is also studied in [12]. The MSE performance for the coherent MAC asymptotically approaches to zero as the number of sensors increases to infinity. However, this is not the case for the orthogonal MAC where the MSE reaches a finite non-zero value as the number of sensor nodes increases [9]. Diversity order of estimation distortion is introduced in [3] and shown to be given by the number of sensors.

In most WSN applications, the source parameter is a physical quantity like temperature, pressure, humidity, sound, ... etc. Therefore, the sensor node observations are correlated where the correlation coefficient is exponentially decaying with distance. In literature, simple signal models were usually assumed. For example, unity correlated observations are assumed in [3], where sensor nodes measure a noisy version of the source signal. The correlation between the observations and the source signal (denoted hereafter as the source-node correlation) is considered in [3]. However, in this model the correlation between observations (denoted hereafter as the inter-node correlation) is determined by the source-node correlation as we will show later. Considering a correlation model with inter-node correlation that determined by the distance between sensor nodes is a more realistic assumption [10], [13].

In this paper, we study the distributed estimation of a scalar parameter where sensor nodes transmit their observations to the FC over a coherent MAC. The observations are spatially correlated and corrupted by noise. Moreover, the communication channel is subject to fading and Additive White Gaussian Noise (AWGN). The FC uses the received signal to estimate the source parameter using LMMSE estimator. The distance-based correlation model of [14] is used in this paper to characterize the source-node correlation and the inter-node correlation. The effect of the distance-based correlation on the estimation performance is demonstrated and compared with the case of unity correlated observations. The outage probability is adopted as the performance measure. A new closed-form expression for the outage probabil-

ity in terms of quadratic forms is introduced. It is shown that less correlated observations degrade the performance.

Hereafter, small letters, bold small letters, and bold capital letters will designate scalars, vectors, and matrices, respectively. If \mathbf{A} is a matrix, then \mathbf{A}^H , \mathbf{A}^T , and $\text{eig}(\mathbf{A})$ denote the hermitian, the transpose, and the eigenvalues of \mathbf{A} , respectively. We define $\text{diag}(\mathbf{a})$ to be a diagonal matrix formed from vector \mathbf{a} . The function $[\cdot]$ returns the same value for the positive values and zero for the negative values and the function $\max(\cdot)$ returns the maximum value.

II. SYSTEM MODEL

Consider a WSN consisting of N sensor nodes and a FC as shown in Fig. 1. Sensor nodes are required to observe a scalar parameter modeled by a zero-mean complex Gaussian random variable $s \sim \mathcal{CN}(0, \sigma_s^2)$. The signals measured by individual sensor nodes can be described as

$$x_i = s_i + n_i, \quad i = 1, \dots, N, \quad (1)$$

where s_i is the i th sensor node observation and $n_i \sim \mathcal{CN}(0, \sigma_n^2)$ is the observation noise. The signals s and $s_i, i = 1, \dots, N$, are modeled as zero-mean joint Gaussian random variables, i.e. $\text{E}\{s_i\} = 0, i = 1, \dots, N$, $\text{E}\{ss_i\} = \rho_i \sigma_s^2, i = 1, \dots, N$, where ρ_i is the source-node correlation coefficient between s and s_i . Moreover, sensor node observations $s_i, i = 1, \dots, N$, are correlated to each other, i.e. $\text{E}\{s_i s_j\} = \rho_{ij} \sigma_s^2, i, j = 1, \dots, N, i \neq j$, where ρ_{ij} is the inter-node correlation coefficient between s_i and s_j .

The correlation coefficients are non-negative and decrease monotonically with distance. Using the power exponential model presented in [14], ρ_i and ρ_{ij} are functions of d_i and d_{ij} , respectively, according to the relation

$$\rho(d) = e^{-\left(\frac{d}{\theta_1}\right)^{\theta_2}}, \theta_1 > 0, 0 < \theta_2 \leq 2, d \in \{d_i, d_{ij}\}, \quad (2)$$

where d_i is the distance between the event source and the i th node, d_{ij} is distance between sensor i and j , $\rho_{ii} = 1$, and θ_1 and θ_2 are the model parameters, where θ_1 normalizes the distance and θ_2 controls the correlation decay rate. Let us define the matrix \mathbf{C} as the cross-node correlation matrix with ρ_{ij} is the element at the i th row and j th column. Assuming random sensor node locations, \mathbf{C} will be a full-rank matrix, and therefore this model will be referred to as *full-rank* correlation model. The aforementioned model for the correlation coefficients ρ_i and ρ_{ij} is more generalized than the one used in [3]–[5] where the signal is given as $x_i(t) = s(t) + v_i(t), i = 1, \dots, N$, and thus $s_i(t) = s(t)$ in this case. This results in *unity* correlated source and observations, i.e. $\rho_i = 1$ and $\rho_{ij} = 1, i, j = 1, \dots, N$. Also, the signal model used in [9], [15] is $x_i(t) = h_i s(t) + v_i(t), i = 1, \dots, N$, which corresponds to $s_i(t) = h_i s(t)$ in Eq. (1) and results in the special case $\rho_{ij} = \rho_i \rho_j$ where $\text{E}\{s(t)s_i(t)\} = \sigma_s^2 h_i = \sigma_s^2 \rho_i$ and $\text{E}\{s_i(t)s_j(t)\} = \sigma_s^2 h_i h_j = \sigma_s^2 \rho_i \rho_j$. Let us define the vector $\mathbf{r} = [\rho_1 \ \rho_2 \ \dots \ \rho_N]^T$, then the cross-node correlation matrix for the later case is given by $\mathbf{C} = \mathbf{r}\mathbf{r}^T$ which is a rank one matrix and thus will be referred to as *rank-one*

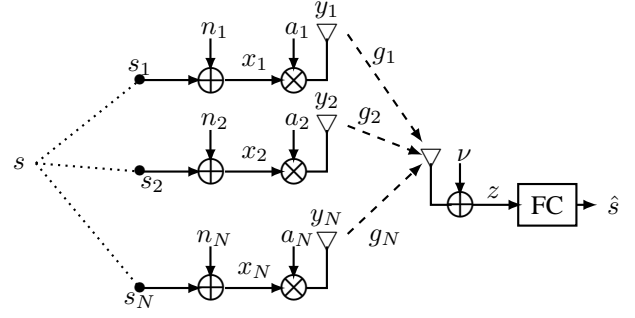


Fig. 1. System model.

correlation model. In the rest of the paper, we will compare between the three correlation models.

For the amplify-and-forward scheme, the transmitted signal from the i th sensor node is given by $y_i = a_i x_i, i = 1, \dots, N$, where a_i is the amplification factor. Here, the transmit power for each node is $\text{E}\{|y_i|^2\} = a_i^2(\sigma_s^2 + \sigma_n^2)$. Assume that the Channel State Information (CSI) is available at the FC. Then, the coherent combining of the transmitted signals received at the FC is

$$z = \sum_{i=1}^N a_i g_i x_i + \nu, \quad (3)$$

where ν is the communication noise $\nu \sim \mathcal{CN}(0, \sigma_\nu^2)$ and g_i is the Rayleigh fading for the i th node $g_i \sim \mathcal{CN}(0, \sigma_g^2)$.

Given the signal and channel statistics, the LMMSE estimate \hat{s} can be expressed as

$$\hat{s} = \frac{\text{E}\{zs\}}{\text{E}\{z^2\}} z = \frac{z \sigma_s^2 \sum_{i=1}^N a_i g_i \rho_i}{\sigma_s^2 \sum_{i=1}^N \sum_{j=1}^N a_i a_j g_i g_j \rho_{ij} + \sum_{i=1}^N \sigma_n^2 a_i^2 g_i^2 + \sigma_\nu^2} \quad (4)$$

and the corresponding distortion becomes

$$D = \text{E}\{(s - \hat{s})^2\} = \sigma_s^2 - \frac{(\text{E}\{zs\})^2}{\text{E}\{z^2\}} = \sigma_s^2 - \frac{\sigma_s^4 \left(\sum_{i=1}^N a_i g_i \rho_i \right)^2}{\sigma_s^2 \sum_{i=1}^N \sum_{j=1}^N a_i a_j g_i g_j \rho_{ij} + \sigma_n^2 \sum_{i=1}^N a_i^2 g_i^2 + \sigma_\nu^2}. \quad (5)$$

The normalized distortion $\tilde{D} = \frac{D}{\sigma_s^2}$ is then given by Eq. (6) (shown at the bottom of the next page). This expression is the generalization of the equivalent one given in [5, Eq. 2] for the unity correlated source and observations. The first term in the numerator and the first term of the denominator in Eq. (6) are the result of the inter-node correlation.

III. EFFECT OF THE CORRELATION ON THE DISTORTION

In this section, the special cases of unity correlation and rank-one model are compared to full-rank model. For simplicity, the channel fading is neglected, i.e. $g_i = 1, i = 1, \dots, N$, and Equal Power Allocation (EPA) is assumed,

i.e. $a_i = \sqrt{P_{\text{tot}}/N(\sigma_s^2 + \sigma_n^2)} = a$, where P_{tot} is the total transmit power for all sensor nodes. Accordingly, the normalized distortion expression for the full-rank model reduces to

$$\tilde{D}_0^{\text{FR}} = \frac{\sigma_s^2 a^2 \left(\sum_{i=1}^N \sum_{j=1}^N \rho_{ij} - \left(\sum_{i=1}^N \rho_i \right)^2 \right) + N \sigma_n^2 a^2 + \sigma_\nu^2}{\sigma_s^2 a^2 \left(\sum_{i=1}^N \sum_{j=1}^N \rho_{ij} \right) + N \sigma_n^2 a^2 + \sigma_\nu^2}. \quad (8)$$

Considering the signal model in [9, Eq. 10], which results in the special case $\rho_{ij} = \rho_i \rho_j$ (rank-one model), the corresponding normalized distortion is given by

$$\tilde{D}_0^{\text{RO}} = \frac{\frac{P_{\text{tot}} \sigma_n^2}{(\sigma_s^2 + \sigma_n^2)} + \sigma_\nu^2}{\frac{P_{\text{tot}} \sigma_s^2}{N(\sigma_s^2 + \sigma_n^2)} \left(\sum_{i=1}^N \rho_i \right)^2 + \frac{P_{\text{tot}} \sigma_n^2}{(\sigma_s^2 + \sigma_n^2)} + \sigma_\nu^2}. \quad (9)$$

Finally, for unity correlated source and observations, i.e. $\rho_i = 1$ and $\rho_{ij} = 1$, $\forall i, j$, the normalized distortion expression reduces to

$$\tilde{D}_0^{\text{U}} = \frac{\frac{P_{\text{tot}} \sigma_n^2}{(\sigma_s^2 + \sigma_n^2)} + \sigma_\nu^2}{\frac{P_{\text{tot}}}{(\sigma_s^2 + \sigma_n^2)} (N \sigma_s^2 + \sigma_n^2) + \sigma_\nu^2}. \quad (10)$$

which is equivalent to the expression in [5, Eq. 2].

Comparing the aforementioned expressions, it is apparent that \tilde{D}_0^{RO} and $\tilde{D}_0^{\text{U}} \rightarrow 0$ as the number of sensor nodes goes to infinity. However, \tilde{D}_0^{FR} does not vanish under the same condition. The effect of correlation on the distortion when the number of sensor nodes increases is depicted in Fig. 2. Eqs. (8), (9), and (10) are averaged over 1000 random realization of sensor node locations and plotted for increasing N . Here, the observation Signal-to-Noise Ratio (SNR) is defined as $\sigma_s^2/\sigma_n^2 = 20$ dB and the communication SNR as $(\sigma_s^2 + \sigma_n^2)/\sigma_\nu^2 = 20$ dB, where $\sigma_s^2 = 1$. The correlation model has $\theta_1 = 250$ and $\theta_2 = 1$. The total transmit power of all sensor nodes is $P_{\text{tot}} = 10$ dB (normalized to σ_s^2). As expected, the distortion for the unity correlated case tends to zero as N increases. The same behavior is noticed for the rank-one model, however with slightly higher distortion in this case. Conversely, the distortion for full-rank model exhibit a floor behavior at a non-zero distortion (≈ 0.182). Therefore, correlation between observations should be taken into consideration when designing distributed estimation schemes because weak correlation degrades its performance.

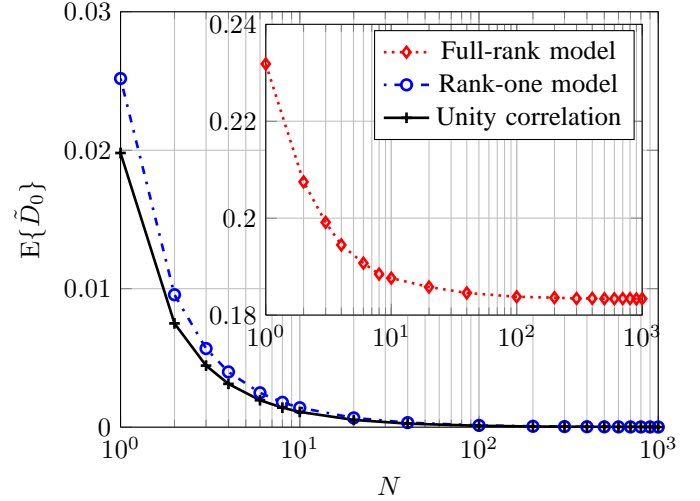


Fig. 2. The behavior of the normalized distortion as the number of sensor nodes N increases for the three correlation models ($\theta_1 = 250$, $\theta_2 = 1$, $\sigma_s^2/\sigma_n^2 = 20$ dB, $\sigma_x^2/\sigma_\nu^2 = 20$ dB, $\sigma_s^2 = 1$, and $P_{\text{tot}} = 10$ dB).

IV. OUTAGE PROBABILITY ANALYSIS

The normalized distortion in Eq. (6) depends on the channel fading, correlation coefficients, channel and measurement noise power, and amplification factors. Given the distance-based correlation model, the correlation coefficients will be invariable for specific network geometry. Therefore, with fixed power allocation and noise power, the channel fading is the only determining factor for the fluctuation in the estimation distortion. In many applications, the interest is to characterize the maximum distortion rather than the average distortion. In such situations, the outage probability is used as a performance measure where the outage refers to the event at which a desired performance level cannot be satisfied. Here, the outage probability is defined as the probability that the normalized distortion exceeds a certain value $\tilde{\delta}$,

$$P_{\text{out}}(\tilde{\delta}) = \Pr(\tilde{D} \geq \tilde{\delta}). \quad (11)$$

To find a closed-form expression for the outage probability, we first express the distortion in vector form. Let us define the vectors $\mathbf{g} = [g_1 \ g_2 \ \dots \ g_N]^T$, $\mathbf{a} = [a_1 \ a_2 \ \dots \ a_N]^T$, and \mathbf{r} and \mathbf{C} as defined previously. It follows that $\sum_{i=1}^N a_i g_i \rho_i = \mathbf{z}^H \mathbf{g}$, where $\mathbf{z} = \mathbf{W} \mathbf{r}$ and $\mathbf{W} = \text{diag}(\mathbf{a})$. Similarly, $\left(\sum_{i=1}^N a_i g_i \rho_i \right)^2 = (\mathbf{z}^H \mathbf{g})^2 = \mathbf{g}^H \mathbf{F} \mathbf{g} = \|\mathbf{g}\|_{\mathbf{F}}^2$, where $\mathbf{F} = \mathbf{z} \mathbf{z}^H$. Also, $\sum_{i=1}^N \sum_{j=1}^N a_i a_j g_i g_j \rho_{ij} =$

$$\tilde{D} = \frac{\sigma_s^2 \left(\sum_{i=1}^N \sum_{j=1}^N a_i a_j g_i g_j \rho_{ij} - \left(\sum_{i=1}^N a_i g_i \rho_i \right)^2 \right) + \sigma_n^2 \sum_{i=1}^N a_i^2 g_i^2 + \sigma_\nu^2}{\sigma_s^2 \sum_{i=1}^N \sum_{j=1}^N a_i a_j g_i g_j \rho_{ij} + \sigma_n^2 \sum_{i=1}^N a_i^2 g_i^2 + \sigma_\nu^2}. \quad (6)$$

$\mathbf{g}^H \mathbf{B} \mathbf{g} = \|\mathbf{g}\|_{\mathbf{B}}^2$, where $\mathbf{B} = \mathbf{W} \mathbf{C} \mathbf{W}^T$ and $\sum_{i=1}^N a_i^2 g_i^2 = \mathbf{g}^H \mathbf{W}^2 \mathbf{g} = \|\mathbf{g}\|_{\mathbf{W}^2}^2$. Then the normalized distortion can be expressed in terms of indefinite quadratic forms as

$$\tilde{D} = \frac{\|\mathbf{g}\|_{\mathbf{B}_1}^2 + \sigma_\nu^2}{\|\mathbf{g}\|_{\mathbf{B}_2}^2 + \sigma_\nu^2}, \quad (12)$$

where $\mathbf{B}_1 = \sigma_s^2 (\mathbf{B} - \mathbf{F}) + \sigma_n^2 \mathbf{W}^2$ and $\mathbf{B}_2 = \sigma_s^2 \mathbf{B} + \sigma_n^2 \mathbf{W}^2$. Accordingly, the outage probability takes the form

$$\begin{aligned} P_{\text{out}}(\tilde{\delta}) &= \Pr \left(\frac{\|\mathbf{g}\|_{\mathbf{B}_1}^2 + \sigma_\nu^2}{\|\mathbf{g}\|_{\mathbf{B}_2}^2 + \sigma_\nu^2} \geq \tilde{\delta} \right) \\ &= \Pr \left(\tilde{\delta} \|\mathbf{g}\|_{\mathbf{B}_2}^2 - \|\mathbf{g}\|_{\mathbf{B}_1}^2 \leq \sigma_\nu^2 - \sigma_\nu^2 \tilde{\delta} \right) \\ &= \Pr \left(\|\mathbf{g}\|_{\mathbf{E}(\tilde{\delta})}^2 \leq (1 - \tilde{\delta}) \sigma_\nu^2 \right), \end{aligned} \quad (13)$$

where $\mathbf{E}(\tilde{\delta}) = \tilde{\delta} \mathbf{B}_2 - \mathbf{B}_1$. Using the results of [16], the outage probability $P_{\text{out}}(\tilde{\delta})$ can be expressed as

$$\begin{aligned} P_{\text{out}}(\tilde{\delta}) &= u \left((1 - \tilde{\delta}) \sigma_\nu^2 \right) + \sum_{l=1}^N \frac{(-\lambda_l)^N}{\prod_{i, l \neq i} (\lambda_i - \lambda_l)} \frac{1}{\lambda_l} \\ &\quad \times e^{-\frac{(1-\tilde{\delta})\sigma_\nu^2}{\lambda_l}} u \left(\frac{(1 - \tilde{\delta}) \sigma_\nu^2}{\lambda_l} \right), \end{aligned} \quad (14)$$

where $\lambda_l = \lambda_l(\tilde{\delta})$, $i = 1, \dots, N$, are the eigenvalues of the matrix $\mathbf{E}(\tilde{\delta})$ and $u(\cdot)$ is the Heaviside unit-step function. This expression can be simplified if \mathbf{E} is substituted by its component matrices,

$$\begin{aligned} \lambda_l(\tilde{\delta}) &= \text{eig} \left(\mathbf{E}(\tilde{\delta}) \right) = \text{eig} \left(\tilde{\delta} \mathbf{B}_2 - \mathbf{B}_1 \right) \\ &= \text{eig} \left(\sigma_s^2 \mathbf{F} - \sigma_s^2 (1 - \tilde{\delta}) \mathbf{B} - \sigma_n^2 (1 - \tilde{\delta}) \mathbf{W}^2 \right) \\ &\approx \sigma_s^2 \text{eig} \left(\mathbf{F} - (1 - \tilde{\delta}) \mathbf{B} \right), \quad l = 1, \dots, N, \end{aligned} \quad (15)$$

where the last approximation results from considering that $\sigma_n^2 \ll \sigma_s^2$. Recall that the matrix \mathbf{F} is the outer product of the vector \mathbf{z} by itself, hence it is of rank one and positive semidefinite. Moreover, \mathbf{B} is also positive semidefinite matrix. Therefore, \mathbf{E} has only one non-negative eigenvalue and all other eigenvalues can simply canceled off in (14) because of the unit-step function. Finally, the expression of outage probability simplifies to

$$P_{\text{out}}(\tilde{\delta}) = 1 - \frac{\lambda_+^{N-1} e^{-\frac{(1-\tilde{\delta})\sigma_\nu^2}{\lambda_+}}}{\prod_{i, l \neq i} (\lambda_+ - \lambda_i)}, \quad 0 \leq \tilde{\delta} \leq 1, \quad (16)$$

where $\lambda_+ = \lambda_+(\tilde{\delta})$ is the only non-negative eigenvalue of $\mathbf{E}(\tilde{\delta})$. Using Weyls inequality [17], this eigenvalue can be lower bounded by

$$\begin{aligned} \lambda_+(\tilde{\delta}) &\approx \sigma_s^2 \max \left(\text{eig} \left(\mathbf{F} - (1 - \tilde{\delta}) \mathbf{B} \right) \right) \\ &\geq \left[\sigma_s^2 \left(\lambda_{\max}^{\mathbf{F}} - (1 - \tilde{\delta}) \lambda_{\max}^{\mathbf{B}} \right) \right], \end{aligned} \quad (17)$$

where $\lambda_{\max}^{\mathbf{F}}$ and $\lambda_{\max}^{\mathbf{B}}$ are the largest eigenvalues of the matrices \mathbf{F} and \mathbf{B} , respectively. The outage probability

depends on the eigenvalues of $\mathbf{E}(\tilde{\delta})$ (exact expression) or the eigenvalues of \mathbf{F} and \mathbf{B} (lower bound).

Assume equal power allocation, one has

$$\begin{aligned} \lambda_+(\tilde{\delta}) &\approx \frac{P_{\text{tot}} \sigma_s^2}{N (\sigma_s^2 + \sigma_n^2)} \max \left(\text{eig} \left(\mathbf{r} \mathbf{r}^H - (1 - \tilde{\delta}) \mathbf{C} \right) \right) \\ &= \frac{P_{\text{tot}} \sigma_s^2}{(\sigma_s^2 + \sigma_n^2)} \tilde{\lambda}_+(\tilde{\delta}) \end{aligned} \quad (18)$$

where $\tilde{\lambda}_+(\tilde{\delta}) = \max \left(\text{eig} \left(\mathbf{r} \mathbf{r}^H - (1 - \tilde{\delta}) \mathbf{C} \right) \right) / N$ is the factor reflecting the effect of the correlation on the largest eigenvalue and it depends on the geometry of the WSN. Considering that the term $P_{\text{tot}} \sigma_s^2 / (\sigma_s^2 + \sigma_n^2)$ is constant for specific network setting, the factor $\tilde{\lambda}_+(\tilde{\delta})$ will be referred to as the normalized eigenvalue. The outage probability in this case is expressed as

$$\begin{aligned} P_{\text{out}}(\tilde{\delta}) &= 1 - \frac{\tilde{\lambda}_+^{N-1}}{\prod_{i, l \neq i} (\tilde{\lambda}_+ - \tilde{\lambda}_i)} \\ &\quad \times \exp \left\{ -\frac{(\sigma_s^2 + \sigma_n^2) (1 - \tilde{\delta}) \sigma_\nu^2}{P_{\text{tot}} \sigma_s^2 \tilde{\lambda}_+} \right\}, \quad 0 \leq \tilde{\delta} \leq 1. \end{aligned} \quad (19)$$

Clearly, the outage probability is a monotonically decreasing function of $\tilde{\lambda}_+(\tilde{\delta})$ and thus related to the signal and observations correlation.

V. NUMERICAL RESULTS

In this section, the analytic results are confirmed by numerical simulations. Consider a WSN that consists of $N = 10$ sensor nodes randomly located in a square area with side length of 20 m. It is required to estimate a source parameter located 30 m away from the center of the sensor nodes. The settings of Fig. 2 are assumed for all the following simulations unless otherwise stated. Moreover, the channel fading has variance $\sigma_g^2 = 1$. All simulation results are averaged over 1000 independent runs.

Fig. 3 shows the outage probability vs. the normalized distortion. The analytic expressions are plotted with solid, dashed, and dotted lines and the simulations are plotted with circle, diamond, and plus marks. The closed-form expression is shown to be in perfect match with the simulation results and the accuracy of the closed-form approximation of the outage probability is verified. The correlation clearly affects the outage performance where more outage occurs for both distance-based correlation models. Note that for the full-rank model, $P_{\text{out}} = 1$ for \tilde{D} with values less than ≈ 0.182 (i.e. the distortion is always larger than this value) which agrees with the results of Fig. 2.

Fig. 4 compares between the exact expression for the largest eigenvalue (normalized by $\frac{P_{\text{tot}} \sigma_s^2}{\sigma_s^2 + \sigma_n^2}$) and its approximation in Eq. (17) for different source locations. The source distance to the center of the square area is set to 50 m, 30 m,

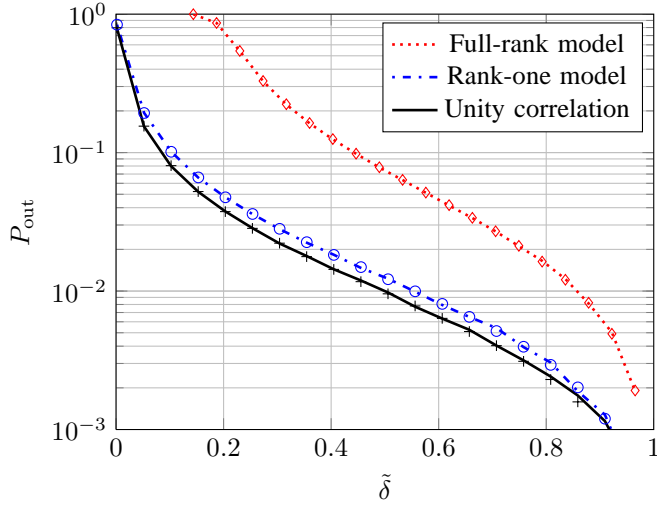


Fig. 3. Outage probability vs. normalized distortion ($\theta_1 = 250$, $\theta_2 = 1$, $\sigma_s^2/\sigma_n^2 = 20$ dB, $\sigma_x^2/\sigma_v^2 = 20$ dB, $\sigma_s^2 = 1$, and $P_{\text{tot}} = 10$ dB).

and 0 m. Note that λ_+ is upper bounded by $\frac{P_{\text{tot}}\sigma_s^2}{\sigma_s^2 + \sigma_n^2}$ since

$$\begin{aligned} \lambda_+(\tilde{\delta}) &\approx \frac{P_{\text{tot}}\sigma_s^2}{N(\sigma_s^2 + \sigma_n^2)} \\ &\times \left(\max(\text{eig}(\mathbf{r}\mathbf{r}^H)) - (1 - \tilde{\delta}) \max(\text{eig}(\mathbf{C})) \right) \\ &\leq \frac{P_{\text{tot}}\sigma_s^2}{N(\sigma_s^2 + \sigma_n^2)} \sum_{i=1}^N \rho_i^2 \leq \frac{P_{\text{tot}}\sigma_s^2}{\sigma_s^2 + \sigma_n^2}. \end{aligned} \quad (20)$$

Recall that the outage probability is a monotonically decreasing function of $\lambda_+(\tilde{\delta})$ and situations with larger eigenvalue corresponds to a better outage performance (i.e. lower outage probability). From the figure, we conclude that observing farther source reduces the largest eigenvalue and thus increases the outage probability.

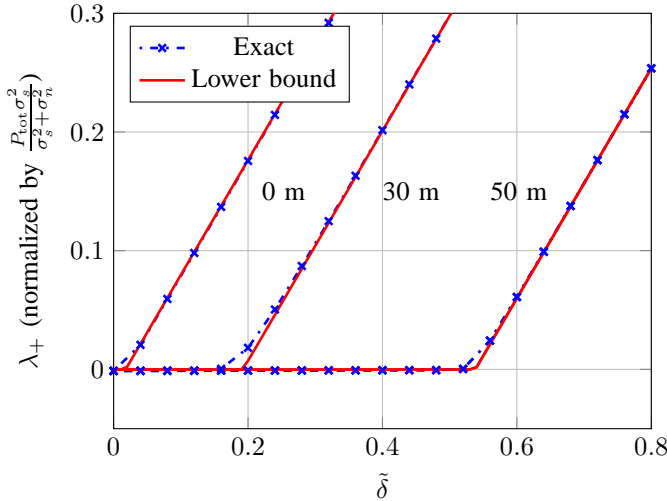


Fig. 4. Largest eigenvalue and its lower bound vs. normalized distortion ($\theta_1 = 250$, $\theta_2 = 1$, $\sigma_s^2/\sigma_n^2 = 20$ dB, $\sigma_x^2/\sigma_v^2 = 20$ dB, $\sigma_s^2 = 1$, and $P_{\text{tot}} = 10$ dB).

VI. CONCLUSIONS

In this paper, distributed estimation of a scalar parameter in WSNs is considered. Correlated source signal and observations are assumed and the effect of correlation is investigated. A closed-form expression for the outage probability is derived to link between the correlation and the outage performance. It is shown that higher distortion levels occurs with higher probability when assuming correlated observations as compared to unity correlated ones. Moreover, the distortion does not vanish when increasing the number of sensor nodes indefinitely for the distance-based correlation.

REFERENCES

- [1] I. Akyildiz, W. Su, Y. Sankarsubramaniam, and E. Cayirci, "Wireless Sensor Networks: A survey," *Computer Networks*, vol. 38, pp. 393–422, Mar. 2002.
- [2] J.-J. Xiao, A. Ribeiro, Z.-Q. Luo, and G. B. Giannakis, "Distributed compression-estimation using wireless sensor networks," *IEEE Sig. Proc. Magazine*, vol. 23, No. 4, pp. 27–41, Jul. 2006.
- [3] S. Cui, J. J. Xiao, A. J. Goldsmith, Z. Q. Luo, and H. V. Poor, "Estimation diversity and energy efficiency in distributed sensing," *IEEE Trans. Sig. Proc.*, vol. 55, no. 9, pp. 4683–4695, Sep. 2007.
- [4] H. Senol and C. Tepedelenlioglu, "Performance of distributed estimation over unknown parallel fading channels," *IEEE Trans. on Sig. Proc.*, vol. 56, no. 12, pp. 6057–6068, Dec. 2008.
- [5] C. H. Wang, A. S. Leong, and S. Dey, "Distortion outage minimization and diversity order analysis for coherent multiaccess," *IEEE Trans. Sig. Proc.*, vol. 59, no. 12, pp. 6144–6159, Dec. 2011.
- [6] T. C. Aysal and K. E. Barner, "Constrained decentralized estimation over noisy channels for sensor networks," *IEEE Trans. on Sig. Proc.*, vol. 56, no. 4, pp. 1398–1410, Apr. 2008.
- [7] X. Wang and C. Yang, "Decentralized estimation over orthogonal multiple-access fading channels in wireless sensor networks: optimal and suboptimal estimators," *EURASIP Journal on Advances in Sig. Proc.*, vol. 2011, no. 1, Dec. 2011.
- [8] I. Bahceci and A. 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.
- [9] J.-J. Xiao, S. Cui, Z.-Q. Luo, and A. J. Goldsmith, "Linear coherent decentralized estimation," *IEEE Trans. on Sig. Proc.*, vol. 56, no. 2, pp. 757–770, Feb. 2008.
- [10] M. H. Chaudhary and L. Vandendorpe, "Power constrained linear estimation in wireless sensor networks with correlated data and digital modulation," *IEEE Trans. on Sig. Proc.*, vol. 60, no. 2, pp. 570–584, Feb. 2012.
- [11] M. Gastpar and M. Vetterli, "Source-channel communication in sensor networks," in *Proc. 2nd International Conference on Information Processing in Sensor Networks (IPSN'03)*, Palo Alto, CA, USA, pp. 162–177, Dec. 2003.
- [12] A. S. Leong and S. Dey, "On scaling laws of diversity schemes in decentralized estimation," *IEEE Trans. Inf. Theory*, vol. 57, no. 7, pp. 4740–4759, Jul. 2011.
- [13] T. Al-Murad, M. Ghogho, and A. Swami, "Effects of fading and spatial correlation on node selection for estimation in wireless sensor networks," in *Proc. IEEE Eleventh International Workshop on Sig. Proc. Advances in Wireless Communications (SPAWC'10)*, Marrakech, Morocco, Jun. 2010.
- [14] M. C. Vuran and I. F. Akyildiz, "Spatial correlation-based collaborative medium access control in wireless sensor networks," *IEEE/ACM Trans. Networking*, vol. 14, no. 2, pp. 316–329, Apr. 2006.
- [15] A. S. Leong, S. Dey, G. N. Nair, and P. Sharma, "Power allocation for outage minimization in state estimation over fading channels," *IEEE Trans. Sig. Proc.*, vol. 59, no. 7, pp. 3382–3397, Jul. 2011.
- [16] T. Y. Al-Naffouri and B. Hassibi, "On the distribution of indefinite quadratic forms in Gaussian random variables," in *Proc. IEEE International Symposium on Information Theory (ISIT'09)*, Seoul, Korea, pp. 1744–1748, Jul. 2009.
- [17] Rajendra Bhatia, *Matrix Analysis*, Graduate Texts in Mathematics, Springer-Verlag, New York, 1997.