This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE Globecom 2010 proceedings.

Optimal Sensor Density in a Distortion-Tolerant Linear Wireless Sensor Network

Jingxian Wu and Ning Sun
Department of Electrical Engineering,
University of Arkansas, Fayetteville, AR 72701, USA.

***Abstract*— The optimum sensor node density in a large linear wireless sensor network with spatial source correlation is studied. Unlike most previous works that rely on the design metric of network capacity with an *error-free* communication assumption, this paper performs analysis under a *distortion-tolerant* communication framework, where controlled distortion in the recovered information is allowed as long as the information can be recovered beyond a certain fidelity. The impacts of node density and spatial data correlation on the information distortion are investigated asymptotically by considering a large network with infinite area, infinite node numbers, but finite node density. Under fixed energy per unit area, it is discovered that: 1) for applications that only need to recover data at discrete locations, placing exact one sensor at the desired measurement locations will generate the optimum performance; 2) for applications that need to recover data at arbitrary locations in the measurement field, the optimum node density is a function of the spatial data correlation.**

I. Introduction

A wireless sensor network (WSN) provides autonomous monitoring of physical or environmental conditions by using a group of spatially distributed sensor nodes transmitting measured data to a fusion center (FC) [1]. Node density, *i.e.*, the number of nodes in a unit area, is a critical design parameter for a WSN. Given a fixed transmission power per unit area, a higher node density means less power per node but more data samples per unit area. Such a trade-off relationship necessitates the study of the optimum node density in a WSN.

There have been considerable works in the literature investigating the impacts of node density on the network performance with various performance metrics [2] - [6], [8]. The seminal work by Gupta and Kumar [2] discovers that the per node throughput in an ad hoc network scales with $\mathcal{O}\left(\frac{1}{\sqrt{N \log N}}\right)$, with N being the number of nodes per unit area. The result in [2] does not consider the spatial data correlation. Data collected in the real world often contain redundancies due to the spatial correlation inherent in the monitored object(s). In [3], a Weiner process is used to model the spatial correlation of a one-dimension field. It is demonstrated that, due to the spatial data correlation, distortion-free communication can be achieved even if the per node throughput tends to 0 as $N \rightarrow \infty$. The optimum node density maximizing the total information under an energy constraint is obtained in [4]. The above study is for peer-to-peer networks, where there are equal numbers of sources and destination. For many-to-one networks such as a WSN, the result is more pessimistic. It is demonstrated in [5] that no compression scheme is sufficient to achieve distortion-free communication in a many-to-one network.

The analysis in most of the previous works is performed by using the design metric of network capacity, which is the maximum throughput supported by a network with *error-free* communication. In reality, a small amount of error might be acceptable for certain applications [7]. In [8], the optimum network density of a many-to-one linear network is studied by using the detection error probability as the performance metric. However, the result is only applicable to a system that performs binary detection at the FC. More sophisticated detections might be required by practical systems.

In this paper, we investigate the optimum node density in a distortion-tolerant linear network with spatial data correlation. In recognition of the distortion-tolerance of many practical applications, we replace the commonly used network capacity metric with a more realistic metric that measures the distortion, *i.e.*, the mean square error (MSE) between the recovered information and the original information. Unlike the binary detection scheme employed in [8], the FC attempts to reconstruct the data from the sensors by exploiting the spatial data correlation with the minimum mean square error (MMSE) receiver. The impacts of node density and spatial data correlation on MSE are investigated asymptotically in a large network with infinite size, infinite nodes, but finite node density, and fixed power per unit area. The analysis accurately captures the trade-off relationship between signal-to-noise ratio (SNR) per node, which is inverse proportional to node density, and spatial sample correlation, which increases with the node density. Optimum node densities for two types of networks, one needs to recover data only at discrete locations, and one needs to recover data at arbitrary locations, are identified through the asymptotic analysis.

II. Problem Formulation

A. System Model

Consider an array of N sensor nodes uniformly placed over a length-L linear section with equal distance d between two adjacent nodes, as shown in Fig. 1. The n-th node is placed at a location with coordinate $l_n = (n-1)d$. The system has a node density of $\delta = \frac{1}{d}$. Each sensor will measure a location dependent physical quantity, $x(l_n)$, such as vibration intensity of a bridge, PH value in soil, etc. Data collected from two points that are close to each other are often strongly correlated due to the spatial redundancy of the measured object. It is assumed that the physical quantities to be measured form a spatial random process that is wide-sense stationary (WSS) in the spatial domain. It is normalized to unit variance with the spatial domain auto-correlation function given by

$$R_{ss}(l_1 - l_2) = \mathbb{E}\left[x(l_1)x(l_2)\right] = \rho^{|l_1 - l_2|d}, \tag{1}$$

where $\rho \in [0, 1]$ is defined as the spatial correlation coefficient, and $\mathbb{E}(\cdot)$ represents mathematical expectation.

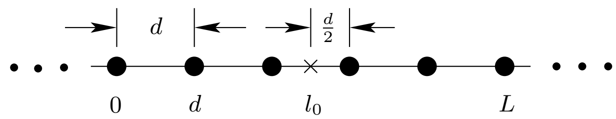

Fig. 1. The linear sensor network

Sensors deliver the measured data to the FC through an orthogonal media access control (MAC) scheme, such as the deterministic frequency division multiplexing access (FDMA), or the random exponentially-interval MAC (EI-MAC) [9], such that collision-free communication is achieved at the FC. The signal observed by the FC from the n-th sensor node is

$$y_n = \sqrt{P_n}x(l_n) + z_n, \tag{2}$$

where P_n is the average transmission power of the n-th node, and z_n is the additive white Gaussian noise (AWGN) with variance σ_z^2. The system model does not consider the effects of fading or pathloss. The subsequent analysis can be easily extended to systems with fading. It is assumed that the total power per unit area is fixed at P_0. Given a linear network with node density δ, the power per node is then $P_n = \frac{P_0}{\delta}$.

B. Optimum MMSE Detection

The FC will obtain an estimate of the spatially continuous quantity, $x(l)$, $\forall l \in [0, L]$, over the entire measurement field, by extracting information transmitted by the N sensors, $\mathbf{y} = [y_1, \cdots, y_N]^T \in \mathcal{R}^{N\times 1}$, where $(\cdot)^T$ represents matrix transpose and $\mathcal{R}$ is the set of real numbers. The MSE at location l is

$$\sigma_l^2 = \mathbb{E}\left[\hat{x}(l) - x(l)\right]^2, l \in [0, L] \tag{3}$$

where $\hat{x}(l)$ is the estimated data at the FC.

The optimum linear receiver that minimizes σ_l^2 is the MMSE receiver described as follows [10]

$$\hat{x}(l) = \sqrt{P_n}\mathbf{r}_l\left(P_n\mathbf{R}_{xx} + \sigma_z^2\mathbf{I}_N\right)^{-1}\mathbf{y}, \tag{4}$$

where $\mathbf{I}_N$ is a size-N identity matrix, $\mathbf{r}_l = \mathbb{E}\left[x(l)\mathbf{x}\right] = [\rho^{|l|}, \rho^{|l-d|}, \cdots, \rho^{|l-Nd|}]^T \in \mathcal{R}^{N\times 1}$, and $\mathbf{R}_{xx} = \mathbb{E}\left[\mathbf{x}\mathbf{x}^T\right] \in \mathcal{R}^{N\times N}$ with the (m, n)-th element being $\rho^{|m-n|d}$.

With the optimum MMSE receiver given in (4), the MSE σ_l^2 can be calculated as [10]

$$\sigma_l^2 = 1 - \mathbf{r}_l^T\left(\mathbf{R}_{xx} + \frac{\delta}{\gamma_0}\mathbf{I}_N\right)^{-1}\mathbf{r}_l, \tag{5}$$

where $\gamma_0 = \frac{P_0}{\sigma_z^2}$ is the SNR per unit area. The MSE σ_l^2 given in (5) is a function of the location l, the SNR γ_0, the spatial correlation coefficient ρ, and the node density δ.

Given a fixed transmission power per unit area, the node density, δ, plays a critical role on the MSE σ_l^2. A smaller node density means more transmission power per node, thus a better SNR per sample. On the other hand, a smaller node density means less samples per unit area, or smaller correlation among the samples, and this will degrade the estimation accuracy.

In order to distinguish the opposite impacts of the node density on the SNR per sample and the spatial sample correlation, we decompose the MMSE receiver described in (4) into two steps as follows.

Definition 1: **Two-Step MMSE**:

1) The FC first obtains an estimate of the data at the sensor locations: $\mathbf{x}_s = [x(l_1), \cdots, x(l_N)]^T \in \mathcal{R}^{N\times 1}$, with a linear MMSE receiver as

$$\hat{\mathbf{x}}_s = \mathbf{W}_s^T\mathbf{y}, \tag{6}$$

where $\hat{\mathbf{x}}_s = [\hat{x}(l_1), \cdots, \hat{x}(l_N)]^T$ is an estimate of $\mathbf{x}_s$. The MMSE matrix $\mathbf{W}_s \in \mathcal{R}^{N\times N}$ is designed to minimize the normalized MSE (NMSE):

$$\sigma_{s,N}^2 = \frac{1}{N}\text{trace}\left[\mathbb{E}(\|\hat{\mathbf{x}}_s - \mathbf{x}_s\|^2)\right], \tag{7}$$

where trace$(\mathbf{A})$ returns the trace of the matrix $\mathbf{A}$, and $\|\mathbf{a}\| = \sqrt{\mathbf{a}\mathbf{a}^T}$ is the L_2-norm of the column vector $\mathbf{a}$.

2) The FC obtains the estimate of a data at an arbitrary location, $\hat{x}(l)$, by interpolating $\hat{\mathbf{x}}_s$ with the MMSE criterion,

$$\hat{x}(l) = \mathbf{w}_{sl}^T\hat{\mathbf{x}}_s, \tag{8}$$

where the vector, $\mathbf{w}_{sl} \in \mathcal{R}^{N\times 1}$, is designed to minimize the MSE $\sigma_l^2 = \mathbb{E}\left[\hat{x}(l) - x(l)\right]^2$.

Lemma 1: The two-step MMSE receiver described in Definition 1 is equivalent to the optimum MMSE given in (4).

Proof: In step 2, the MMSE vector $\mathbf{w}_{sl}$ that minimizes σ_l^2 can be obtained through the orthogonal principal [10], $\mathbb{E}\{[\mathbf{w}_{sl}^T\hat{\mathbf{x}}_s - x(l)]\hat{\mathbf{x}}_s^T\} = 0$, and the result is

$$\begin{aligned} \mathbf{w}_{sl}^T &= \mathbb{E}[x(l)\hat{\mathbf{x}}_s^T]\left[\mathbb{E}(\hat{\mathbf{x}}_s\hat{\mathbf{x}}_s^T)\right]^{-1}, \\ &= \sqrt{P_n}\mathbf{r}_l\left(P_n\mathbf{R}_{xx} + \sigma_z^2\mathbf{I}_N\right)^{-1}\mathbf{W}_s^{-1}, \end{aligned} \tag{9}$$

where (2) and (6) are used for the second equality. Combining (6), (8), and (9) leads to (4). ∎

Decomposing the optimum MMSE of (4) into the two-step MMSE in Definition 1 allows us to study the two opposite effects of the node density on the MSE separately. We investigate, respectively, the impacts of the node density on the performance of the two steps in the next two sections.

III. MMSE Estimation at Sensor Locations

The asymptotic behavior of the MMSE estimation of the data at the sensor locations as described in (6) is discussed in this section.

A. Asymptotic Analysis

For the MMSE estimation described in (6), the optimum $\mathbf{W}_s$ that minimizes the NMSE, $\sigma_{s,N}^2$, can be found through the orthogonal principal, $\mathbb{E}\left[(\hat{\mathbf{x}}_d - \mathbf{x}_s)\mathbf{y}\right] = 0$. The result is

$$\mathbf{W}_s^T = \sqrt{P_n}\mathbf{R}_{xx}\left(P_n\mathbf{R}_{xx} + \sigma_z^2\mathbf{I}\right)^{-1}. \tag{10}$$

The error correlation matrix, $\mathbf{R}_{ee} = \mathbb{E}\left[\mathbf{e}_s\mathbf{e}_s^T\right]$, with $\mathbf{e}_s = \hat{\mathbf{x}}_s - \mathbf{x}_s$, can then be calculated as

$$\mathbf{R}_{ee}=\mathbf{R}_{xx}-\mathbf{R}_{xx}\left(\mathbf{R}_{xx}+\frac{\delta}{\gamma_0}\mathbf{I}\right)^{-1}\mathbf{R}_{xx}=\left(\mathbf{R}_{xx}^{-1}+\frac{\gamma_0}{\delta}\mathbf{I}_N\right)^{-1}, \quad (11)$$

where the orthogonal principal is used in the first equality, and the second equality is based on the identify $\mathbf{D}^{-1}+\mathbf{D}^{-1}\mathbf{C}(\mathbf{A}-\mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1}\mathbf{B}\mathbf{D}^{-1} = (\mathbf{D}-\mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1}$.

The NMSE can then be calculated as $\sigma_{s,N}^2 = \frac{1}{N}\text{trace}(\mathbf{R}_{ee})$. From (11), the calculation of the NMSE involves matrix inversion and the trace operation. In order to explicit identify the impacts of node density and spatial data correlation on the NMSE, we resort to the asymptotic analysis by letting $N \to \infty$ while keeping a finite node density δ. The results are presented as follows.

Proposition 1: When $N \to \infty$ while keeping a finite δ, the NMSE of the estimated data at the sensor locations is

$$\sigma_s^2 = \lim_{N\to\infty}\sigma_{s,N}^2 = \left[\left(1+\frac{\gamma_0}{\delta}\right)^2 + \frac{4\gamma_0\rho^{\frac{2}{\delta}}}{\delta\left(1-\rho^{\frac{2}{\delta}}\right)}\right]^{-\frac{1}{2}}. \quad (12)$$

Proof: Performing the eigenvalue decomposition of $\mathbf{R}_{xx}$ in (11), we can rewrite the NMSE as

$$\sigma_{s,N}^2 = \frac{1}{N}\sum_{n=1}^{N}\left(\frac{1}{\lambda_n}+\frac{\gamma_0}{\delta}\right)^{-1}, \quad (13)$$

where λ_n is the n-th eigenvalue of $\mathbf{R}_{xx}$. Based on Szego's Theorem [11], when $N \to \infty$, (13) can be rewritten as

$$\sigma_s^2 = \lim_{N\to\infty}\sigma_{s,N}^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi}\left[\frac{1}{\Lambda(\omega)}+\frac{\gamma_0}{\delta}\right]^{-1}d\omega, \quad (14)$$

where $\Lambda(\omega) = \sum_{n=-\infty}^{+\infty}\rho^{|n|d}e^{-jn\omega}$ is the discrete-time Fourier transform (DTFT) of the sequence, $\left\{\rho^{|n|d}\right\}_n$, which are elements of the Toeplitz matrix $\mathbf{R}_{xx}$. The DTFT, $\Lambda(\omega)$, can be calculated by

$$\Lambda(\omega) = \frac{1-\rho^{2d}}{1+\rho^{2d}-2\rho^d\cos\omega}. \quad (15)$$

Substituting (15) into (14) leads to

$$\sigma_s^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi}\left[\frac{1+\rho^{2d}-2\rho^d\cos\omega}{1-\rho^{2d}}+\frac{\gamma_0}{\delta}\right]^{-1}d\omega. \quad (16)$$

The above integral can be solved by using the identity [12, eqn. (2.553.3)]

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}(a+b\cos\omega)^{-1}d\omega = \frac{1}{\sqrt{a^2-b^2}}, \quad (17)$$

Combining (16) with (17) leads to (12). ■

The opposite effects of node density on the NMSE are manifested in (12), where the asymptotic NMSE is explicitly expressed as a function of the node density and the spatial correlation. From (12), the node density affects the NMSE in the form of two functions, $f_1(\delta) = \frac{\gamma_0}{\delta}$, and $f_2(\delta) = \frac{\rho^{\frac{2}{\delta}}}{1-\rho^{\frac{2}{\delta}}}$.

The asymptotic NMSE, σ_s^2, increases as either of the two functions decreases. The first function, $f_1(\delta)$, can be interpreted as the SNR per node, which is inverse proportional to δ. Thus $f_1(\delta)$ translates a positive correlation between δ and the NMSE. The second function, $f_2(\delta)$, is related to the spatial correlation among sensors, and it is an increasing function of δ. Hence, $f_2(\delta)$ translates a negative correlation between δ and the NMSE. As analyzed above, δ exhibits two opposite effects on NMSE through $f_1(\delta)$ and $f_2(\delta)$. For the estimation of data at the sensor node locations, it is shown in the following corollary that the effects of the SNR per node, $f_1(\delta)$, dominates that of the spatial correlation, $f_2(\delta)$.

Corollary 1: The asymptotic NMSE given in (12) is a monotonic increasing function of the node density, δ.

Proof: From (12), it is equivalent to show that $g_1(d) = (1+\gamma_0 d)^2 + 4\gamma_0 d\frac{\rho^{2d}}{1-\rho^{2d}}$ is a monotonic increasing function of $d = \frac{1}{\delta}$. Taking the first derivative of $g_1(d)$, we have

$$g_1'(d) = \frac{2\gamma_0}{\left(1-\rho^{2d}\right)^2}\times g_2(d,\gamma_0), \quad (18)$$

where $g_2(d,\gamma_0)$ is defined as

$$g_2(d,\gamma_0) \triangleq \left(1-\rho^{2d}\right)^2(1+\gamma_0 d)+2\rho^{2d}\left(1-\rho^{2d}\right)+4d\log(\rho)\rho^{2d} \quad (19)$$

From (18), in order to prove $g_1'(d) \geq 0$, it is sufficient to prove that $g_2(d,0) \geq 0$ because $g_2(d,\gamma_0) \geq g_2(d,0)$. Let $v = \rho^{2d} \in [0,1]$, then $g_2(d,0)$ can be rewritten as

$$g_3(v) \triangleq g_2(d,0) = 1 - v^2 + 2v\log(v), \quad 0 \leq v \leq 1 \quad (20)$$

It can be easily shown that $g_3''(v)=2(\frac{1}{v}-1) \geq 0, \forall v \in [0,1]$, therefore $g_3(v)$ is quadratic on $[0,1]$ with the minimum value obtained at the solution of $g_3'(v) = -2v + 2\log(v) + 2 = 0$, which is $v = 1$. Substituting $v = 1$ into (20), we have $\min\{g_3(v)\} = 0$. Therefore, $g_2(d,\gamma_0) \geq g_2(d,0) = g_3(v) \geq 0$, and this completes the proof. ■

B. Optimum Node Density

The result in Corollary 1 indicates that, the NMSE for estimating data at the sensor node locations can benefit from a smaller density. *Therefore, if we only want to obtain the data at some discrete locations, we should use a node density that is as small as allowed by the application, i.e., placing exactly one sensor at each desired measurement location will obtain the optimum performance.*

Fig. 2 shows the asymptotic NMSE as a function of the node density, δ, under various values of the spatial correlation coefficient, ρ. The SNR per unit area is $\gamma_0 = 10$ dB. The simulation result is obtained by using $N = 1,000$ nodes to approximate infinite nodes. In the simulation, data samples are assumed to be a Gaussian process with zero mean and spatial auto-correlation function given in (1). Good agreement is observed between the asymptotic analytical results and the simulation results. As pointed out by Corollary 1, the NMSE increases monotonically as δ increases, indicating the dominance of the SNR per node over the spatial correlation.

In addition, it can be seen from Fig. 2 that the NMSE approaches a constant value as $\delta \to \infty$. This indicates a

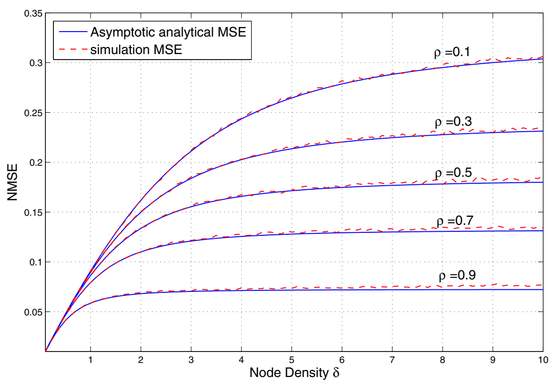

Fig. 2. The asymptotic NMSE as a function of the node density δ.

balance between the opposite effects between $f_1(\delta)$ and $f_2(\delta)$ as $\delta \to \infty$. This is corroborated by the following corollary.

Corollary 2: For the estimation of the data at the sensor locations, the asymptotic NMSE is upper bounded by

$$\sigma_s^2 \leq \left[1 - \frac{2\gamma_0}{\log(\rho)}\right]^{-\frac{1}{2}} \tag{21}$$

Proof: Eqn. (21) can be directly proved by setting $\lim_{\delta\to\infty} \sigma_s^2$ in (12). ■

The NMSE upper bound is determined by the spatial correlation and the SNR per unit area. The result in Corollary 2 indicates that when δ becomes large enough, the impact of the SNR per node is balanced by that of the spatial correlation.

IV. MMSE Spatial Interpolation

The impacts of the node density on the spatial interpolation as described in (8) are investigated in this section.

A. Asymptotic Analysis

Once the estimates of the data at the sensor locations are obtained, they can be interpolated to obtain an estimate of the data at an arbitrary location in the measurement field.

As discussed in Definition 1 and Lemma 1, MMSE spatial interpolation can obtain the optimum performance, with the minimum MSE σ_l^2 given in (3). The MSE given in (3) depends on the location l. Since we are interested in the reconstruction fidelity of the entire measurement field, here we will consider the worst case scenario by estimating the data located in the middle between two consecutive sensors, with coordinate $l_0 = l_k + \frac{d}{2}$, as shown in Fig. 1, when $N \to \infty$.

It is difficult to express σ_l^2 in (3) as an explicit function of δ. To explicitly investigate the impact of the node density on the reconstruction fidelity, we will resort to a sub-optimum MMSE algorithm by only using the estimated data from the sensor that is the closest to the measurement point as

$$\hat{x}(l_0) = w_{l_0}\hat{x}(l_k), \tag{22}$$

where $\hat{x}(l_k)$ is the k-th component of $\hat{\mathbf{x}}_s$ estimated from (6). Since the spatial correlation among the sensors has been taken into consideration during the estimation of $\hat{x}(l_k)$, it is expected that estimating $x(l_0)$ from its closet sensor, $\hat{x}(l_k)$, will have a similar performance as estimating $x(l_0)$ from all the sensors, $\hat{\mathbf{x}}_s$. This statement is verified through numerical examples in the next subsection.

Based on the orthogonal principal, $\mathbb{E}\{[w_{l_0}\hat{x}(l_k) - x(l_0)]\,\hat{x}(l_k)\} = 0$, the value of w_{l_0} that can minimize the MSE is $w_{l_0} = r_{x\hat{x}} \cdot r_{\hat{x}\hat{x}}^{-1}$, where $r_{x\hat{x}} = \mathbb{E}[x(l_0)\hat{x}(l_k)]$, and $r_{\hat{x}\hat{x}} = \mathbb{E}[\hat{x}(l_k)\hat{x}(l_k)]$. The MSE, $\sigma_{l_0}^2 = \mathbb{E}\left[(w_{l_0}\hat{x}(l_k) - x(l_0))\right]^2$, can be calculated by

$$\sigma_{l_0}^2 = 1 - r_{x\hat{x}}^2 r_{\hat{x}\hat{x}}^{-1}. \tag{23}$$

Denote $e(l_k) = \hat{x}(l_k) - x(l_k)$, then $r_{x\hat{x}}$ can be expressed as

$$r_{x\hat{x}} = \mathbb{E}[x(l_0)x(l_k)] + \mathbb{E}[x(l_0)e(l_k)] \approx \rho^{\frac{d}{2}}(1 + r_{xe}), \tag{24}$$

where $r_{xe} = \mathbb{E}[x(l_k)e(l_k)]$, and the approximation $\mathbb{E}[x(l_0)e(l_k)] \approx \mathbb{E}[x(l_0)x(l_k)]r_{xe}$ is used in the derivation. The approximation makes intuitive sense because it calculates $\mathbb{E}[x(l_0)e(l_k)]$ by scaling $\mathbb{E}[x(l_k)e(l_k)]$ with $\rho^{\frac{d}{2}}$ to account for the $\frac{d}{2}$ distance between l_k and l_0. Numerical examples show that the approximation is reasonably accurate.

The terms r_{xe} and $r_{\hat{x}\hat{x}}$ are related to step one of the estimation, and they can be asymptotically calculated as

$$r_{xe} = \lim_{N\to\infty} \frac{1}{N}\text{trace}\left(\mathbf{R}_{se}\right), \tag{25a}$$

$$r_{\hat{x}\hat{x}} = \lim_{N\to\infty} \frac{1}{N}\text{trace}\left(\mathbf{R}_{\hat{x}\hat{x}}\right), \tag{25b}$$

where $\mathbf{R}_{se} = \mathbb{E}\left[\mathbf{x}_s\mathbf{e}_s^T\right]$, and $\mathbf{R}_{\hat{x}\hat{x}} = \mathbb{E}\left[\hat{\mathbf{x}}_s\hat{\mathbf{x}}_s^T\right]$. Based on the orthogonal principal, it is easy to show that $\mathbf{R}_{\hat{x}\hat{x}}=\mathbf{R}_{xx}-\mathbf{R}_{ee}$ and $\mathbf{R}_{se} = -\mathbf{R}_{ee}$. Therefore, we have $r_{xe} = -\sigma_s^2$, and $r_{\hat{x}\hat{x}} = 1 - \sigma_s^2$, where σ_s^2 is the NMSE in Proposition 1. The identity $\lim_{N\to\infty}\frac{1}{N}\text{trace}\left(\mathbf{R}_{xx}\right) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\Lambda(\omega)d\omega = 1$, which is derived from Szego's Theorem [11], is used in the derivation, with $\Lambda(\omega)$ being the DTFT of $\rho^{|n|d}$ defined in (15).

Combining the above results with (23) and (24), we have an approximation of $\sigma_{l_0}^2$ as

$$\hat{\sigma}_{l_0}^2 = 1 - \rho^{\frac{1}{\delta}}(1 - \sigma_s^2). \tag{26}$$

Since both σ_s^2 and $\rho^{\frac{1}{\delta}}$ are monotonic increasing functions of δ, it is easy to see that $\hat{\sigma}_{l_0}^2$ is a monotonic decreasing function in δ. This can be intuitively explained by the fact that spatial interpolation depends mainly on the spatial correlation among the data samples, and a higher density means a stronger correlation among the data samples.

B. Optimum Node Density

The result in (26) is plotted in Fig. 3 as a function of δ under various values of ρ. In order to verify the accuracy of the approximation, also shown in the figure are the exact optimum MSE obtained from (5). It can be seen from the figure that (26) gives a reasonably good approximation of the exact optimum MSE, and the approximation becomes more accurate as ρ increases.

In addition, it can be seen from the figure that, when δ is small, the MSE decreases dramatically as δ increases. When δ reaches a certain threshold, no apparent performance gain can be achieved by increasing δ further, *i.e.*, the slope of $\hat{\sigma}_{l_0}^2$

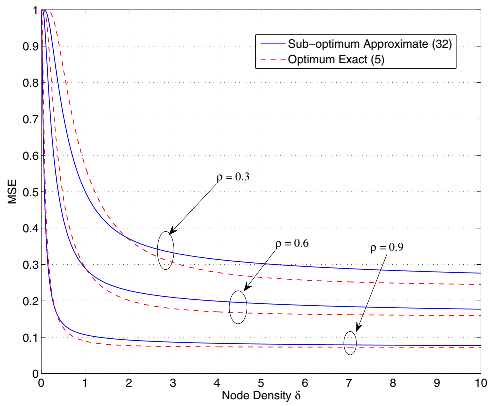

Fig. 3. Comparison of the exact optimum MSE with the approximated sub-optimum MSE

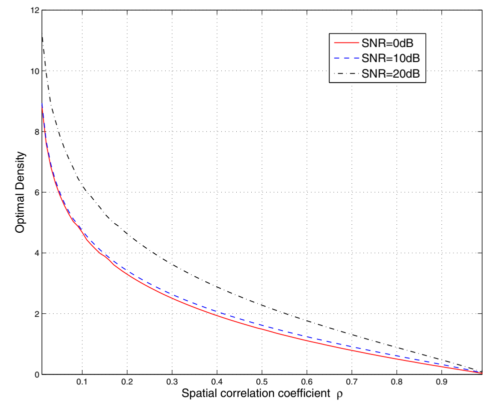

Fig. 4. Optimum node density v.s. spatial correlation coefficient

approaches zero as δ increases. Therefore, the optimum node density can be chosen as the point such that $\left|\frac{\partial \sigma_{l_0}^2}{\partial \delta}\right| \leq \epsilon$, with ϵ being a small number.

Fig. 4 shows the optimum node density as a function of the spatial correlation coefficient. The optimum node density is obtained by numerically solving the equation $\left|\frac{\partial \sigma_{l_0}^2}{\partial \delta}\right|_{\delta_0} = \epsilon$, with $\epsilon = 10^{-5}$. The results in this figure demonstrate that the optimum node density decreases as ρ increases. *Therefore, a smaller density is required for a field with stronger spatial correlation*.

The MSE in (26) is lower bounded as stated in the following corollary.

Corollary 3: The following relationship holds for σ_s^2 and $\hat{\sigma}_{l_0}^2$

$$\hat{\sigma}_{l_0}^2 \geq \left(1 - \frac{2\gamma_0}{\log(\rho)}\right)^{-\frac{1}{2}} \geq \sigma_s^2 \tag{27}$$

Proof: Since $\hat{\sigma}_{l_0}^2$ is a decreasing function of δ, its minimum value can be obtained by letting $\delta \to \infty$ in (26), and (27) follows immediately. ■

The result in (27) indicates that $\hat{\sigma}_{l_0}^2$ is always bigger than σ_s^2 and they converge when $\delta \to \infty$. This can be explained by the fact that the estimation of $x(l_0)$ is based on $\hat{x}(l_k)$, thus the fidelity of $\hat{x}(l_0)$ can not exceed that of $\hat{x}(l_k)$. This result further corroborates that, for the estimation of data at a discrete location, a sensor node needs to be placed at the desired location to ensure the optimum performance.

V. Conclusions

In this paper, the optimum sensor node density for a large linear distortion-tolerant WSN with spatial source correlation was studied. The impacts of the node density on the MSE of the data reconstructed at the FC were investigated through asymptotic analysis, where the number of nodes tends to infinity while the node density remains constant. There were two observations from the analytical results. First, if the network only needs to estimate spatially discrete data, placing exactly one sensor at the desired measurement locations will generate the optimum performance. Second, for the estimation of the data at arbitrary locations in the measurement field, the optimum node density can be found when the MSE-density slope is close to zero, and the optimum density decreases as the spatial correlation coefficient increases.

Acknowledgments

This work was supported in part by the National Science Foundation under Grant ECCS-0917041 and the Arkansas Bioscience Institute.

References

[1] D. Culler, D. Estrin, and M. Srivastava, "Overview of sensor networks," *IEEE Computer*,vol. 37, pp. 41-49, August 2004.

[2] P. Gupta, and P. R. Kumar, "The capacity of wireless networks," *IEEE Trans. on Inform. Theory*, vol. 46, pp. 388-404, Mar. 2000.

[3] S. D. Servetto, "On the feasibility of large scale wireless sensor networks," in *Proc. IEEE Int. Conf. commu., control, and computing*, Oct. 2002.

[4] Y. Sung, H. V. Poor, and H. Yu, "Optimal node density for two-dimensional sensor arrays," in *Proc. IEEE Sensor Array Multichannel Signal Processing Workshop*, pp. 271-274, 2008.

[5] D. Marco, E.J. Duarte-Melo, M. Liu, and D. L. Neuhoff, "On the many-to-one transport capacity of a dense wireless sensor network and the compressibility of its data," *Inform. proc. in sensor networks*, vol. 2634, pp. 1-16, April 2003.

[6] R. Cristescu, and M. Vetterli, "On the optimal density for real-time data gathering of spatio-temporal processes in sensor networks," *IEEE Trans. on Signal Processing*, vol. 51, pp. 159-164, August 2003.

[7] H. Luo, A. Pandya, and G. Pottie, "Detection fidelity in distributed wireless sensor neworks," *Center for Embedded Network Sensing (CENS) Technical Report 20*, July 2003.

[8] J. F. Chamberland, and V. V. Veeravalli, "How dense should a sensor network be for detection with correlated observations?" *IEEE Trans. Information thepry*, vol. 52, pp. 5099-5106, Nov. 2006.

[9] J. Wu, "Ultra-low power compressive wireless sensing for distributed wireless networks," in *Proc. IEEE Mil. Commun. Conf. MILCOM'09*, pp. 1-7, Oct. 2009.

[10] S. M. Kay, *Fundamentals of statistical signal processing, vol. I, Estimation Theory*, NJ: Prentice-Hall, 1993.

[11] H. Gazzah, P. A. Regalia, and J.P. Delmas, "Asymptotic eigenvalue distribution of block Toeplitz matrices and application to blind SIMO channel identification," *IEEE Trans. Inform. Theory*, vol. 47, pp. 1243-1251, Mar. 2001.

[12] I. S. Gradshteyn, and I. M. Ryzhik , *Table of integrals, series, and products*, 6nd ed. Academic Press, pp. 168, July 2000.