

# Optimum Sensor Density in Distortion-Tolerant Wireless Sensor Networks

Jingxian Wu, *Member, IEEE*, and Ning Sun, *Student Member, IEEE*

**Abstract**—The optimum sensor node density for one- and two-dimensional (1-D and 2-D) wireless sensor networks (WSNs) with spatial source correlation is studied in this paper. The WSN attempts to reconstruct a spatially correlated signal field by collecting the location-dependent measurements from the distributed sensor nodes. The WSN is designed to minimize the mean square error (MSE) distortion between the original and the reconstructed signals under the constraint of a fixed power per unit area. The impacts of node density and spatial data correlation on the network performance are investigated for both small networks with finite number of nodes, and large networks with infinite area, infinite number of nodes, but finite node density through asymptotic analysis. The interactions among the various network parameters and their impacts on the system performance are quantitatively identified with exact analytical expressions, many of which are in closed-forms. The results provide guidelines on the design of practical WSNs.

**Index Terms**—Node density, distortion-tolerant communication, wireless sensor network, asymptotic analysis, MMSE.

## 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]. One of the primary challenges faced by the design of a large WSN is to determine the node density, *i.e.*, the number of nodes in a unit area, to optimize the network performance under the energy and/or cost constraints [1]–[12]. Given a fixed transmission power per unit area, a higher node density means less power available to each node, which degrades the network performance due to the reduced signal-to-noise ratio (SNR) for the signal transmitted by each node. On the other hand, a higher node density can obtain more data samples per unit area, which can benefit the system performance. Such a trade-off relationship necessitates the study of the optimum node density in practical WSNs.

There have been considerable works in the literature investigating the impacts of node density on the network performance for both one-dimensional (1-D) and two-dimensional (2-D) networks [1]–[15]. The seminal work by Gupta and Kumar [13] 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, *i.e.*, the node density. The result in [13] 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 [14], a Wiener process is used to model the spatial correlation of an 1-D 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 above study is for peer-to-peer networks, where there are equal numbers of sources and destinations. For many-to-one networks such as a WSN, it is demonstrated in [15] that no compression scheme is sufficient to achieve distortion-free communications.

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* communications. In reality, a small amount of errors might be acceptable for real world applications such as target detection [2], [3], [7], target localization [4], information coverage [5], [6], and information recovery [8]–[12], etc. In [7], the optimum network density of an 1-D network is studied by minimizing the detection error probability. In [8], an arbitrary point on a continuous measurement field is estimated by interpolating the samples collected by the spatially discrete sensors. The studies in [8] only consider the distortion from the spatial interpolation, and the distortions introduced by the noisy channel are not incorporated in the analysis. A distortion lower bound is derived in [10] for a network with a finite number of correlated sources as a function of the number of sensors and spatial-temporal communication bandwidth. The analysis in [10] is only applicable to a measurement field with finite degree-of-freedom.

In [11], the asymptotic optimum sensor density that can maximize the information collected from a 2-D correlated random field is obtained under a total energy constraint as the number of nodes and the measurement field tend to infinity. A total energy constraint in a large network often leads to negligible energy per node or per unit area, yet in reality the energy budget per unit area is non-trivial. A constraint on a fixed power per unit area, or equivalently, fixed energy per sample per unit area, is adopted in our previous work [12] to obtain the asymptotically optimum node density as the number of nodes and the size of the measurement field tend to infinity. The power or energy per sample constraint will limit the instantaneous power of a sensor node, and it can be easily translated to the energy constraint of a node under a fixed sampling rate. The results in [12] are only applicable to 1-D networks, and the analysis is based on an approximation of the

Manuscript received January 19, 2011; revised June 24 and November 4, 2011; accepted December 27, 2011. The associate editor coordinating the review of this paper and approving it for publication was X. Wang.

Part of this work was presented in the IEEE Global Telecommunication Conference 2010, Miami, USA, Dec. 2010. This work was supported in part by the National Science Foundation under Grant ECCS-0917041.

The authors are with the Department of Electrical Engineering, University of Arkansas, Fayetteville, AR 72701, USA (e-mails: {wuj, nsun}@uark.edu). Digital Object Identifier 10.1109/TWC.2012.041612.110127

minimum mean square error (MMSE) spatial interpolation.

In this paper, we investigate the optimum node density in the 1-D and 2-D distortion-tolerant networks under the constraint of a fixed power per unit area. The WSN attempts to reconstruct a spatially correlated signal field by collecting measurements from distributed sensor nodes. There is no assumption on the statistical properties of the field, other than that it forms a continuous random process that is wide sense stationary (WSS) in the space domain. In recognition of the distortion-tolerance of many practical applications, the optimization is performed with respect to the mean square error (MSE) distortion between the original and the recovered signals. Each sensor node collects spatially correlated samples of the measurement field, and forwards the information to a data FC through a noisy link. The FC reconstructs an estimate of the signal field by exploiting the spatial data correlation with an optimum MMSE receiver. Under the constraint of a fixed transmission power per unit area, the impacts of node density and spatial data correlation on MSE are investigated for both small networks with finite number of nodes and large networks with infinite area but finite node density. Exact analytical expressions are obtained to describe the interactions and tradeoff relationship between the SNR per node, which is inversely proportional to the node density, and spatial sample correlation, which increases with the node density, for both 1-D and 2-D networks. It is observed that the 1-D and 2-D networks have similar performance trends and their performance difference diminishes as the spatial correlation increases.

The remainder of this paper is organized as follows. Section II introduces the system model and a two-step MMSE estimation method. Sections III and IV study the impacts of the node density on the performance of 1-D and 2-D networks, respectively. In these two sections, the optimum node densities in various networks are identified, and numerical examples are presented to demonstrate the interactions among the various system parameters. Section V concludes the paper.

## II. PROBLEM FORMULATION

### A. System Model

Consider a WSN with  $N$  sensor nodes evenly distributed over a measurement field  $\Omega_\eta$ , as shown in Fig. 1(a) for an 1-D network, and Fig. 1(b) for a 2-D network. The smallest distance between two adjacent sensors is  $d$ . The  $n$ -th node is placed at a location with coordinate  $\eta_n \in \Omega_\eta$ , for  $n = 1, 2, \dots, N$ . The uniform node distribution shown in Fig. 1(a) is adopted here for mathematical tractability, and similar configurations have been widely used in the literature [7]–[9], [11], [12], and [14]. Define the node density,  $\delta$ , as the number of nodes in a unit area.

Each sensor will measure a location dependent physical quantity,  $x(\eta_n)$ , such as vibration intensity of a bridge, and humidity, etc. Data collected from two sensors close to each other are often strongly correlated due to the spatial redundancy of the measured object. In this paper, the correlation between two data samples are modeled as follows

$$\mathbb{E}[x(\eta_p)x(\eta_q)] = \rho^{\|\eta_p - \eta_q\|}, \quad (1)$$

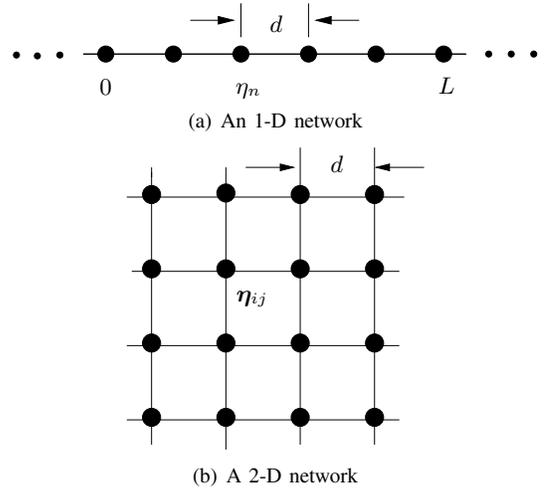


Fig. 1. The 1D and 2D networks with evenly distributed nodes.

where  $\rho \in [0, 1]$  is the spatial correlation coefficient,  $\mathbb{E}(\cdot)$  represents mathematical expectation, and  $\|\mathbf{a}\| = \sqrt{\mathbf{a}\mathbf{a}^T}$  is the  $L_2$ -norm of the column vector  $\mathbf{a}$  with  $(\cdot)^T$  representing matrix transpose.

It is assumed that sensors deliver the measured data to the FC through an orthogonal media access control (MAC) scheme such that collision-free communication is achieved at the FC. The signal observed by the FC from the  $n$ -th sensor node is

$$y_n = \sqrt{P_n}x(\eta_n) + z_n, \quad (2)$$

where  $P_n$  is the average transmission power, or average energy per sample, of the  $n$ -th node, and  $z_n$  is the additive white Gaussian noise (AWGN) with variance  $\sigma_z^2$ . It is assumed that the total power per unit area is fixed at  $P_0$  to ensure a fair comparison among networks with different node densities and different sizes. Given a network with node density  $\delta$ , 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,  $x(\eta)$ ,  $\forall \eta \in \Omega_\eta$ , over the entire measurement field, by extracting information transmitted from the  $N$  sensors,  $\mathbf{y} = [y_1, \dots, y_N]^T \in \mathcal{R}^{N \times 1}$ , where  $\mathcal{R}$  is the set of real numbers. The MSE at any location  $\eta$  is

$$\sigma_\eta^2 = \mathbb{E}[\hat{x}(\eta) - x(\eta)]^2, \quad \eta \in \Omega_\eta \quad (3)$$

where  $\hat{x}(\eta)$  is an estimate of  $x(\eta)$  at the FC.

The optimum linear receiver that minimizes  $\sigma_\eta^2$  is the MMSE receiver described as follows [18]

$$\hat{x}(\eta) = \sqrt{P_n} \mathbf{r}_\eta^T (P_n \mathbf{R}_{ss} + \sigma_z^2 \mathbf{I}_N)^{-1} \mathbf{y}, \quad (4)$$

where  $\mathbf{r}_\eta = \mathbb{E}[x(\eta) \mathbf{x}_s] = [\rho^{\|\eta - \eta_1\|}, \dots, \rho^{\|\eta - \eta_N\|}]^T \in \mathcal{R}^{N \times 1}$  with  $\mathbf{x}_s = [x(\eta_1), \dots, x(\eta_N)]^T$ ,  $\mathbf{R}_{ss} = \mathbb{E}[\mathbf{x}_s \mathbf{x}_s^T] \in \mathcal{R}^{N \times N}$  with the  $(p, q)$ -th element being  $\rho^{\|\eta_p - \eta_q\|}$  as defined in (1), and  $\mathbf{I}_N$  is a size- $N$  identity matrix.

With the optimum MMSE receiver given in (4), the MSE  $\sigma_\eta^2$  can be calculated as [18]

$$\sigma_\eta^2 = 1 - \mathbf{r}_\eta^T \left( \mathbf{R}_{ss} + \frac{\delta}{\gamma_0} \mathbf{I}_N \right)^{-1} \mathbf{r}_\eta, \quad (5)$$

where  $\gamma_0 = \frac{P_0}{\sigma_z^2}$  is the SNR per unit area. The MSE  $\sigma_\eta^2$  given in (5) is a function of the location  $\boldsymbol{\eta}$ , the SNR  $\gamma_0$ , the spatial correlation coefficient  $\rho$ , and the node density  $\delta$ .

Given a fixed transmission power per unit area, the node density,  $\delta$ , plays a critical role on the MSE  $\sigma_\eta^2$ . 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 a 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:  $\mathbf{x}_s = [x(\boldsymbol{\eta}_1), \dots, x(\boldsymbol{\eta}_N)]^T \in \mathcal{R}^{N \times 1}$ , with a linear MMSE receiver as

$$\hat{\mathbf{x}}_s = \mathbf{W}_s^T \mathbf{y}, \quad (6)$$

where  $\hat{\mathbf{x}}_s = [\hat{x}(\boldsymbol{\eta}_1), \dots, \hat{x}(\boldsymbol{\eta}_N)]^T$  is an estimate of  $\mathbf{x}_s$ . The MMSE matrix  $\mathbf{W}_s \in \mathcal{R}^{N \times N}$  is designed to minimize the MSE per node  $\sigma_{s,N}^2 = \frac{1}{N} \mathbb{E} [\|\hat{\mathbf{x}}_s - \mathbf{x}_s\|^2]$ .

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

$$\hat{x}(\boldsymbol{\eta}) = \mathbf{w}_{sl}^T \hat{\mathbf{x}}_s, \quad (7)$$

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

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

*Proof:* The proof is in Appendix A. ■

Decomposing the optimum MMSE of (4) into the two-step MMSE allows us to study the two opposite effects of the node density on the MSE separately. In the next two sections, we will investigate, respectively, the impacts of the node density on the 1-D and 2-D networks by following the two-step MMSE.

### III. OPTIMUM NODE DENSITY IN 1-D NETWORKS

In this section, we study the optimum node density in an 1-D network, where the  $N$  sensor nodes are evenly distributed over a length- $L$  linear section as shown in Fig. 1(a). In a linear network, the  $n$ -th node is placed at a location with the coordinate  $\boldsymbol{\eta}_n = \eta_n = (n-1)d$ . Following the two-step MMSE given in Definition 1, we will study in the next two subsections the impacts of the node density on the performance of the two steps: MMSE estimation of the data at the sensor locations, and MMSE interpolation for the estimation of the data at arbitrary locations.

#### A. MMSE Estimation at Sensor Locations

From (6), the optimum  $\mathbf{W}_s$  that minimizes  $\sigma_{s,N}^2$  can be found through the orthogonal principal,  $\mathbb{E} [(\hat{\mathbf{x}}_s - \mathbf{x}_s) \mathbf{y}^T] = 0$ . The result is

$$\mathbf{W}_s^T = \sqrt{P_n} \mathbf{R}_{ss} (P_n \mathbf{R}_{ss} + \sigma_z^2 \mathbf{I}_N)^{-1}. \quad (8)$$

The error correlation matrix,  $\mathbf{R}_{ee}^{(s)} = \mathbb{E} [\mathbf{e}_s \mathbf{e}_s^T]$ , with  $\mathbf{e}_s = \hat{\mathbf{x}}_s - \mathbf{x}_s$ , can be calculated by

$$\mathbf{R}_{ee}^{(s)} = \mathbf{R}_{ss} - \mathbf{R}_{ss} \left( \mathbf{R}_{ss} + \frac{\delta}{\gamma_0} \mathbf{I}_N \right)^{-1} \mathbf{R}_{ss} = \left( \mathbf{R}_{ss}^{-1} + \frac{\gamma_0}{\delta} \mathbf{I}_N \right)^{-1}, \quad (9)$$

where the orthogonal principal is used in the first equality, and the second equality is based on the identity  $\mathbf{D}^{-1} + \mathbf{D}^{-1} \mathbf{C} (\mathbf{A} - \mathbf{B} \mathbf{D}^{-1} \mathbf{C})^{-1} \mathbf{B} \mathbf{D}^{-1} = (\mathbf{D} - \mathbf{C} \mathbf{A}^{-1} \mathbf{B})^{-1}$ . The MSE can then be calculated as  $\sigma_{s,N}^2 = \frac{1}{N} \text{trace}(\mathbf{R}_{ee}^{(s)})$ , where  $\text{trace}(\mathbf{A})$  returns the trace of the matrix  $\mathbf{A}$ . The calculation of the MSE involves matrix inversion and the trace operation. Performing the eigenvalue decomposition of  $\mathbf{R}_{ss}$  in (9), we have

$$\sigma_{s,N}^2 = \frac{1}{N} \sum_{n=1}^N \left( \frac{1}{\lambda_n} + \frac{\gamma_0}{\delta} \right)^{-1}, \quad (10)$$

where  $\lambda_n$  is the  $n$ -th eigenvalue of  $\mathbf{R}_{ss}$ .

In order to explicitly identify the impacts of the node density and the spatial data correlation on the MSE, we resort to the asymptotic analysis by letting  $N \rightarrow \infty$  while keeping a finite node density  $\delta$ . The results are presented as follows.

*Proposition 1:* When  $N \rightarrow \infty$  while keeping a finite  $\delta$ , the asymptotic MSE of the estimated data at the sensor locations in an 1-D network is

$$\sigma_s^2 \triangleq \lim_{N \rightarrow \infty} \sigma_{s,N}^2 = \left[ \left( 1 + \frac{\gamma_0}{\delta} \right)^2 + \frac{4\gamma_0 \rho^{\frac{2}{\alpha}}}{\delta (1 - \rho^{\frac{2}{\alpha}})} \right]^{-\frac{1}{2}}. \quad (11)$$

*Proof:* The proof is given in Appendix B. ■

In (11), the opposite effects of node density on the asymptotic MSE are manifested in the form of two functions,  $f_1(\delta) \triangleq \frac{\gamma_0}{\delta}$ , and  $f_2(\delta) \triangleq \frac{\rho^{\frac{2}{\alpha}}}{1 - \rho^{\frac{2}{\alpha}}}$ . The first function,  $f_1(\delta)$ , is the SNR per node, which is inversely proportional to  $\delta$ . Thus  $f_1(\delta)$  translates a positive correlation between  $\delta$  and the asymptotic MSE. The second function,  $f_2(\delta)$ , is related to the spatial correlation among sensors, and it is an increasing function of  $\delta$ . Hence,  $f_2(\delta)$  translates a negative correlation between  $\delta$  and the asymptotic MSE. Therefore,  $\delta$  exhibits two opposite effects on MSE 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 effect of the SNR per node,  $f_1(\delta)$ , dominates that of the spatial correlation,  $f_2(\delta)$ .

*Corollary 1:* The asymptotic MSE given in (11) is a monotonic increasing function of the node density,  $\delta$ .

*Proof:* The proof is in Appendix C. ■

The result in Corollary 1 indicates that, the asymptotic MSE 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 MSE as a function of the node density,  $\delta$ . The SNR per unit area is  $\gamma_0 = 10$  dB. Data samples are assumed to be a zero-mean Gaussian process with the spatial auto-correlation function given in (1). The simulation results are obtained by using  $N = 1,000$  nodes to approximate infinite nodes. Excellent agreement is observed

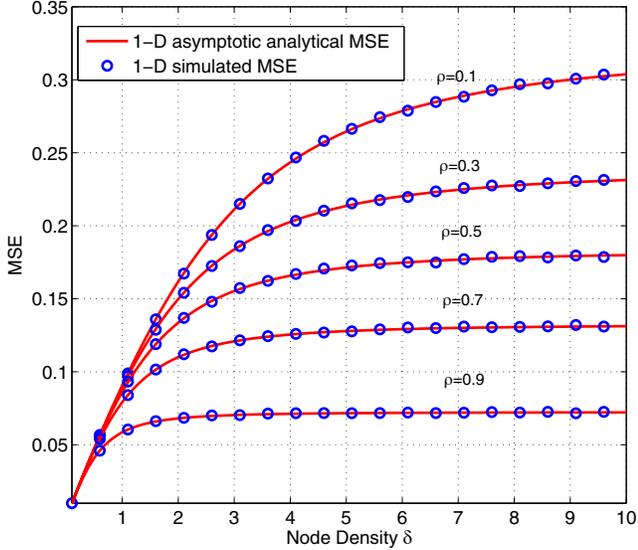


Fig. 2. The asymptotic MSE of the data samples in an 1-D network ( $\gamma_0 = 10$  dB).

between the asymptotic analytical results with  $N \rightarrow \infty$  and the simulation results with  $N = 1,000$ . As pointed out by Corollary 1, the MSE increases monotonically as  $\delta$  increases, indicating the dominance of the SNR per node over the spatial correlation. It can be seen from Fig. 2 that the MSE approaches a constant value as  $\delta \rightarrow \infty$ . This indicates a balance between the opposite effects between  $f_1(\delta)$  and  $f_2(\delta)$  as  $\delta \rightarrow \infty$ , which is corroborated by the following corollary.

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

$$\sigma_s^2 \leq \left(1 - \frac{2\gamma_0}{\log \rho}\right)^{-\frac{1}{2}} \quad (12)$$

*Proof:* Eqn. (12) can be directly proved by setting  $\lim_{\delta \rightarrow \infty} \sigma_s^2$  in (11). ■

The asymptotic MSE upper bound is determined by the spatial correlation and the SNR per unit area.

### B. MMSE Spatial Interpolation

The estimates of the data at the sensor locations can be interpolated to get an estimate of the data at any location.

As discussed in Definition 1 and Lemma 1, MMSE spatial interpolation can obtain the optimum performance by minimizing  $\sigma_\eta^2$  given in (3). The MSE given in (3) depends on the location  $\eta$ . Since we are interested in the reconstruction fidelity of the entire measurement field, the worst case scenario will be considered by estimating the data located in the middle of two neighboring sensors, with coordinate  $\eta'_n = (n - \frac{1}{2})d$ . Correspondingly, denote the data vector to be estimated through interpolation as  $\mathbf{x}_d = [x(\eta'_1), \dots, x(\eta'_N)]^T \in \mathcal{R}^{N \times 1}$ .

Following the orthogonal principal,  $\mathbb{E}[(\hat{\mathbf{x}}_d - \mathbf{x}_d)\hat{\mathbf{x}}_s^T] = \mathbf{0}$ , where  $\hat{\mathbf{x}}_d$  is an estimate of  $\mathbf{x}_d$ , the MMSE spatial interpolation can be expressed by

$$\hat{\mathbf{x}}_d = \mathbf{R}_{d\hat{s}} \mathbf{R}_{\hat{s}\hat{s}}^{-1} \hat{\mathbf{x}}_s, \quad (13)$$

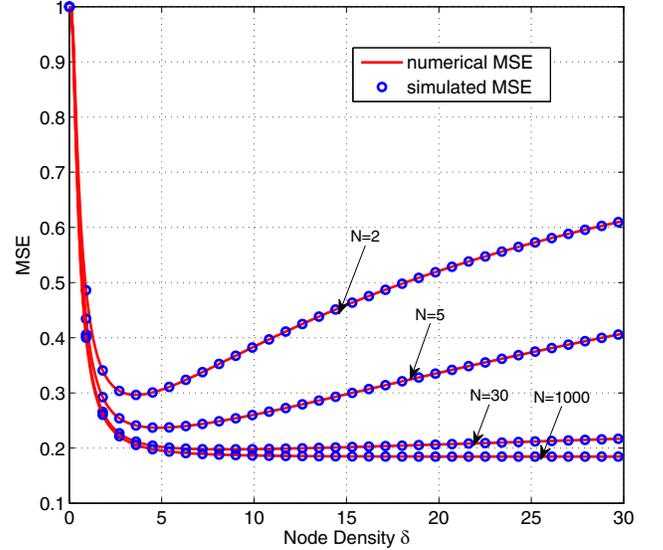


Fig. 3. The MSE of the spatial interpolations under various node number  $N$  in an 1-D network ( $\rho = 0.5$ ,  $\gamma_0 = 10$  dB).

where

$$\mathbf{R}_{d\hat{s}} \triangleq \mathbb{E}(\mathbf{x}_d \hat{\mathbf{x}}_s^T) = \sqrt{P_n} \mathbf{R}_{ds} \mathbf{W}_s, \quad (14a)$$

$$\mathbf{R}_{\hat{s}\hat{s}} \triangleq \mathbb{E}(\hat{\mathbf{x}}_s \hat{\mathbf{x}}_s^T) = \mathbf{W}_s^T (P_n \mathbf{R}_{ss} + \sigma_z^2 \mathbf{I}) \mathbf{W}_s, \quad (14b)$$

with  $\mathbf{R}_{ds} \triangleq \mathbb{E}(\mathbf{x}_d \hat{\mathbf{x}}_s)$  being a Toeplitz matrix. The first row and column of  $\mathbf{R}_{ds}$  are  $[1, 1, \rho^d, \dots, \rho^{(N-2)d}]^T$  and  $[1, \rho^d, \dots, \rho^{(N-1)d}]^T$ , respectively.

Combining (13) with (14), we have

$$\hat{\mathbf{x}}_d = \sqrt{P_n} \mathbf{R}_{ds} (P_n \mathbf{R}_{ss} + \sigma_z^2 \mathbf{I})^{-1} \mathbf{y}. \quad (15)$$

The corresponding error correlation matrix,  $\mathbf{R}_{ee}^{(d)} \triangleq \mathbb{E}[(\hat{\mathbf{x}}_d - \mathbf{x}_d)(\hat{\mathbf{x}}_d - \mathbf{x}_d)^T]$ , can then be calculated by

$$\mathbf{R}_{ee}^{(d)} = \mathbf{R}_{ss} - \mathbf{R}_{ds} \left( \mathbf{R}_{ss} + \frac{\delta}{\gamma_0} \mathbf{I}_N \right)^{-1} \mathbf{R}_{sd}, \quad (16)$$

where  $\mathbf{R}_{dd} = \mathbb{E}(\mathbf{x}_d \mathbf{x}_d^T) = \mathbf{R}_{ss}$  is used in the above equation, and  $\mathbf{R}_{sd} = \mathbf{R}_{ds}^T$ . The MSE for spatial interpolation is  $\sigma_{d,N}^2 = \frac{1}{N} \text{trace}(\mathbf{R}_{ee}^{(d)})$ .

The MSE for the spatial interpolation is numerically evaluated for different values of  $\delta$  and  $N$ , and the results are shown in Fig. 3. The spatial correlation coefficient is  $\rho = 0.5$  and the SNR per unit area is  $\gamma_0 = 10$  dB. When  $N$  is small, e.g.,  $N < 30$ , the MSE is convex in  $\delta$ . When  $N$  is large, e.g.,  $N \geq 30$ , the spatial interpolation MSE becomes a monotonically decreasing function of  $\delta$ . In addition, increasing  $N$  from 30 to 1000 only leads to a marginal performance improvement. The following proposition presents the asymptotic MSE as  $N \rightarrow \infty$ .

*Proposition 2:* When  $N \rightarrow \infty$  while keeping a finite  $\delta$ , the MSE of the MMSE interpolation in an 1-D network is

$$\sigma_d^2 \triangleq \lim_{N \rightarrow \infty} \sigma_{d,N}^2 = \left( \frac{\delta}{\gamma_0} + \frac{1 - \rho^{\frac{1}{\delta}}}{1 + \rho^{\frac{1}{\delta}}} \right)^{\frac{1}{2}} \left( \frac{\delta}{\gamma_0} + \frac{1 + \rho^{\frac{1}{\delta}}}{1 - \rho^{\frac{1}{\delta}}} \right)^{-\frac{1}{2}}. \quad (17)$$

*Proof:* The proof is in Appendix D. ■

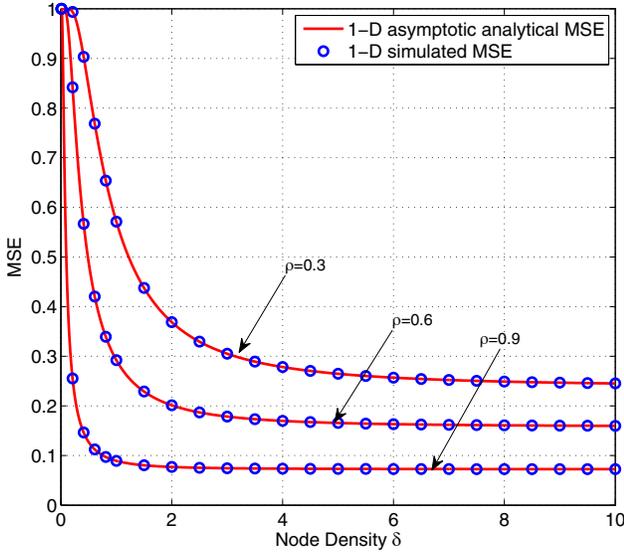


Fig. 4. The asymptotic MSE of the spatial interpolation in an 1-D network ( $\gamma_0 = 10$  dB).

Define  $f_3(\delta) \triangleq \frac{1+\rho^{\frac{1}{\delta}}}{1-\rho^{\frac{1}{\delta}}}$ . It can be easily shown that  $f_3(\delta)$  is an increasing function of  $\delta$ , and its impact on the asymptotic MSE is opposite to the SNR per node,  $f_1(\delta)$ . The following corollary shows that  $\sigma_d^2$  is dominated by the effects of  $f_3(\delta)$ .

*Corollary 3:* The asymptotic MSE given in (17) is a monotonically decreasing function of the node density,  $\delta$ .

*Proof:* The proof is in Appendix E. ■

From Corollaries 1 and 3, it is apparent that  $\delta$  has opposite impacts on  $\sigma_s^2$  and  $\sigma_d^2$ . The results in Corollary 3 can be intuitively explained by the fact that the spatial interpolation depends mainly on the spatial correlation among the data samples, and a higher density means a stronger correlation among the data samples, thus a better estimation fidelity.

The asymptotic MSE of the data interpolation is shown in Fig. 4, where it is apparent that  $\sigma_d^2$  is a decreasing function of  $\delta$ . The simulation parameters are the same as those in Fig. 2. Again, perfect agreement is observed between the simulation results with  $N = 1,000$  and asymptotic analytic results with  $N \rightarrow \infty$ . When  $\delta \rightarrow \infty$ ,  $\sigma_d^2$  is lower bounded, and this is described in the following corollary.

*Corollary 4:* The following inequality holds for  $\sigma_s^2$  and  $\sigma_d^2$

$$\sigma_d^2 \geq \left(1 - \frac{2\gamma_0}{\log \rho}\right)^{-\frac{1}{2}} \geq \sigma_s^2 \quad (18)$$

*Proof:* Since  $\sigma_d^2$  is a decreasing function of  $\delta$ , its minimum value can be obtained by letting  $\delta \rightarrow \infty$  in (17), and (18) follows immediately. ■

The result in (18) indicates that  $\sigma_d^2$  is always bigger than  $\sigma_s^2$  and they converge when  $\delta \rightarrow \infty$ . This can be explained by the fact that the estimation of  $\mathbf{x}_d$  is based on the estimation accuracy of  $\mathbf{x}_s$ , thus the fidelity of  $\hat{\mathbf{x}}_d$  can not exceed that of  $\hat{\mathbf{x}}_s$ . *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, because interpolation will always lead to an inferior performance.*

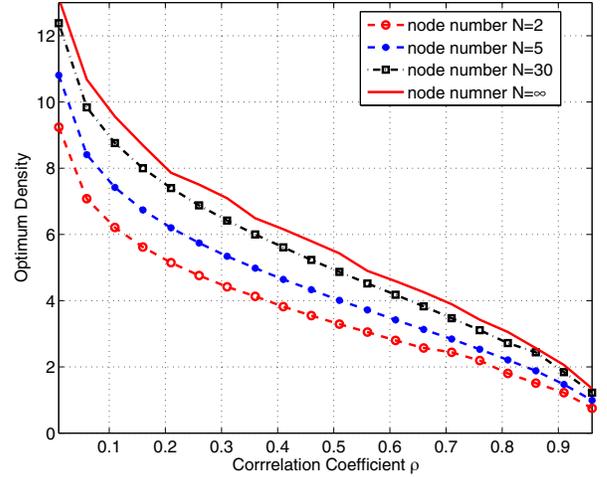


Fig. 5. Optimum node density v.s. correlation coefficient  $\rho$  in an 1-D network ( $\gamma_0 = 10$  dB).

It is observed from Fig. 4 that, when  $\delta$  is small, the MSE decreases dramatically as  $\delta$  increases. When  $\delta$  reaches a certain threshold, no apparent performance gain can be achieved by increasing  $\delta$  further, *i.e.*, the slope of  $\sigma_d^2$  approaches zero as  $\delta$  increases. Therefore, the optimum node density can be chosen as the point such that  $\left|\frac{\partial \sigma_d^2}{\partial \delta}\right| \leq \epsilon$ , with  $\epsilon$  being a small number.

From (17), the slope of  $\sigma_d^2$  can be calculated as

$$\frac{\partial \sigma_d^2}{\partial \delta} = \frac{\sigma_d^2}{2} \cdot \left[ \frac{\frac{1}{\gamma_0} + \frac{2 \log \rho \cdot \rho^{\frac{1}{\delta}}}{\delta^2 \cdot (1+\rho^{\frac{1}{\delta}})^2}}{\frac{\delta}{\gamma_0} + \frac{1-\rho^{\frac{1}{\delta}}}{1+\rho^{\frac{1}{\delta}}}} - \frac{\frac{1}{\gamma_0} - \frac{2 \log \rho \cdot \rho^{\frac{1}{\delta}}}{\delta^2 \cdot (1-\rho^{\frac{1}{\delta}})^2}}{\frac{\delta}{\gamma_0} + \frac{1+\rho^{\frac{1}{\delta}}}{1-\rho^{\frac{1}{\delta}}}} \right] \quad (19)$$

The optimum node density can then be obtained by numerically solving the equation  $\left|\frac{\partial \sigma_d^2}{\partial \delta}\right|_{\delta_0} = \epsilon$ .

Fig. 5 shows the optimum node density as functions of the spatial correlation coefficient  $\rho$  and the number of nodes  $N$ . The SNR per unit area is  $\gamma_0 = 10$  dB. When  $N$  is small ( $N = 2$  and  $N = 5$ ), the MSE is convex in  $\delta$  as shown in Fig. 3, and the optimum node density is obtained by finding the value of  $\delta$  that minimizes the MSE. When  $N$  is large ( $N = 30$  and  $N \rightarrow \infty$ ), the optimum node density is obtained by solving  $\left|\frac{\partial \sigma_d^2}{\partial \delta}\right| \leq \epsilon$  with  $\epsilon = 10^{-3}$ . The optimum node density decreases as  $\rho$  increases for all the systems. In addition, given  $\delta$ , the optimum node density increases as  $N$  increases, and it is upper bounded by the asymptotic case with  $N \rightarrow \infty$ . The optimum node density for  $N = 30$  is very close to the asymptotic results with  $N \rightarrow \infty$ .

#### IV. OPTIMUM NODE DENSITY IN 2-D NETWORKS

The impacts of node density on the estimation fidelity in a 2-D network are studied in this section. In a 2-D network, consider the  $N$  sensor nodes located over a grid on a square region with area  $\sqrt{Nd} \times \sqrt{Nd}$  as shown in Fig. 1(b). The coordinates for the nodes are  $\boldsymbol{\eta}_{ij} = [id, jd]^T$ , for  $i, j = 0, \dots, K-1$  with  $K = \sqrt{N}$ . It should be noted that the node density in a 2-D network is  $\delta = \frac{1}{d^2}$ , which is different from the 1-D case.

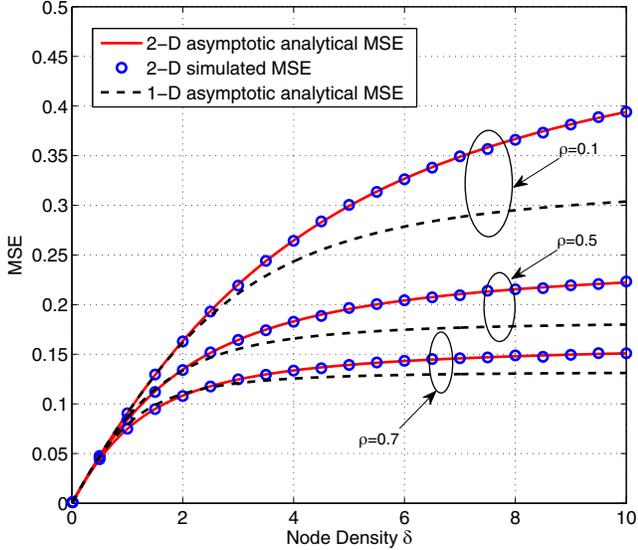


Fig. 6. The asymptotic MSE of the data samples estimation in the 1-D and 2-D networks ( $\gamma_0 = 10$  dB).

Stacking the data from all the sensors into a column vector, we have  $\xi_s = [x_0^T, \dots, x_{K-1}^T]^T \in \mathcal{R}^{N \times 1}$ , where  $x_m = [x(\eta_{m0}), \dots, x(\eta_{m(K-1)})]^T \in \mathcal{R}^{K \times 1}$ . The auto-correlation matrix,  $\Phi_{ss} = \mathbb{E}[\xi_s \xi_s^T] \in \mathcal{R}^{N \times N}$ , can be expressed in the form of a block Toeplitz matrix as

$$\Phi_{ss} = \begin{pmatrix} \mathbf{R}_0 & \mathbf{R}_1 & \cdots & \mathbf{R}_{K-1} \\ \mathbf{R}_1 & \mathbf{R}_0 & \cdots & \mathbf{R}_{K-2} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{R}_{K-1} & \mathbf{R}_{K-2} & \cdots & \mathbf{R}_0 \end{pmatrix}, \quad (20)$$

where  $\mathbf{R}_m = \mathbb{E}(x_m x_0) \in \mathcal{R}^{K \times K}$  is a symmetric Toeplitz matrix with the first row and first column being  $[r_{m,0}, r_{m,1}, \dots, r_{m,K-1}]^T \in \mathcal{R}^{K \times 1}$ , and

$$r_{m,k} = \mathbb{E}[x(\eta_{mk})x(\eta_{00})] = \rho \sqrt{\frac{k^2+m^2}{\delta}}. \quad (21)$$

The matrix,  $\Phi_{ss}$ , assumes the form of a Toeplitz-block-Toeplitz (TBT) matrix [17], i.e.,  $\Phi_{ss}$  is a block Toeplitz matrix, and each sub-matrix is also a Toeplitz matrix.

#### A. MMSE Estimation at Sensor Locations

Following the same procedure as in Section III-A, we have the MSE,  $\vartheta_{s,N}^2 \triangleq \frac{1}{N} \mathbb{E}[\|\hat{\xi}_s - \xi_s\|^2]$ , as

$$\vartheta_{s,N}^2 = \frac{1}{N} \text{trace} \left( \Phi_{ss}^{-1} + \frac{\gamma_0}{\delta} \mathbf{I}_N \right)^{-1}, \quad (22)$$

where  $\hat{\xi}_s$  is the MMSE estimate of  $\xi_s$ . Performing the eigenvalue decomposition of  $\Phi_{ss}$  in (20), we can rewrite the MSE in (22) as

$$\vartheta_{s,N}^2 = \frac{1}{N} \sum_{k=1}^{K-1} \sum_{m=1}^{K-1} \left( \frac{1}{\lambda_{k,m}} + \frac{\gamma_0}{\delta} \right)^{-1}, \quad (23)$$

where  $\lambda_{k,m}$ , for  $k, m = 0, 1, \dots, K-1$ , are eigenvalues of  $\Phi_{ss}$ .

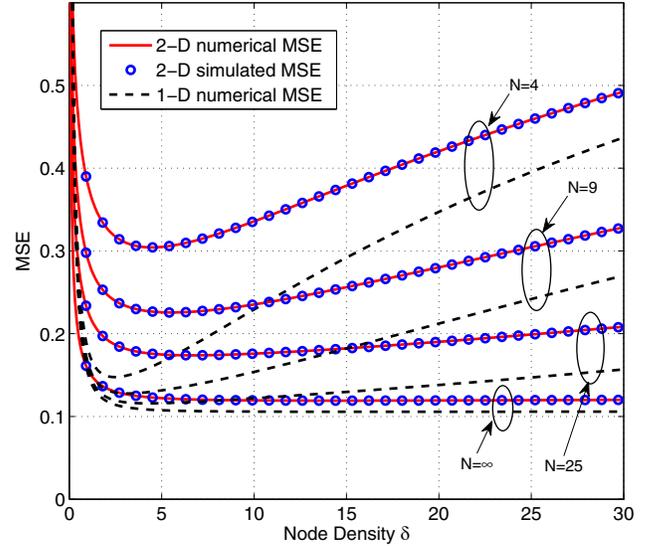


Fig. 7. The MSE of the spatial interpolations under various node number  $N$  in the 1-D and 2-D networks ( $\rho = 0.8$ ,  $\gamma_0 = 10$  dB).

*Proposition 3:* When  $N \rightarrow \infty$  while keeping a finite node density  $\delta$ , the asymptotic MSE of the estimated data at the sensor locations in a 2-D network with spatial correlation coefficient  $\rho$  is

$$\vartheta_s^2 \triangleq \lim_{N \rightarrow \infty} \vartheta_{s,N}^2 = \int_{-\frac{1}{2}}^{\frac{1}{2}} \int_{-\frac{1}{2}}^{\frac{1}{2}} \left[ \frac{1}{\Lambda'_{ss}(f_1, f_2)} + \frac{\gamma_0}{\delta} \right]^{-1} df_1 df_2, \quad (24)$$

where

$$\Lambda'_{ss}(f_1, f_2) = \sum_{k=-\infty}^{+\infty} \sum_{m=-\infty}^{+\infty} \rho \sqrt{\frac{k^2+m^2}{\delta}} e^{-j2\pi(kf_1+mf_2)}. \quad (25)$$

*Proof:* The results in (24) can be obtained by applying [17, Theorem 1] to (23), which is the extension of the Szego's theorem to TBT matrices. ■

The expression in (24) can be easily evaluated numerically given that the integrals are of finite limits. Even though  $\Lambda'_{ss}(f_1, f_2)$  is expressed as the sum of an infinite series, the value of  $\rho \sqrt{\frac{k^2+m^2}{\delta}}$  decreases exponentially as  $k$  and  $m$  increase, thus  $\Lambda'_{ss}(f_1, f_2)$  can be accurately approximated with moderate limits on  $k$  and  $m$ .

Fig. 6 shows the asymptotic MSE  $\vartheta_s^2$  as a function of the node density  $\delta$ . For comparison, the MSE  $\sigma_s^2$  for an 1-D network is also shown in this figure. The SNR is  $\gamma_0 = 10$  dB. Similar to the 1-D case, the asymptotic MSE  $\vartheta_s^2$  is a monotonic increasing function of  $\delta$ . In Fig. 6, given a fixed node density, the asymptotic 2-D MSE is larger (worse) than the asymptotic 1-D MSE. This can be explained by the fact that each node in the 2-D network needs to cover a larger area than its 1-D counterpart. The difference between the 1-D and 2-D cases gradually diminishes as the spatial correlation coefficient,  $\rho$ , increases, because the impact of node distance decreases as  $\rho$  getting close to 1.

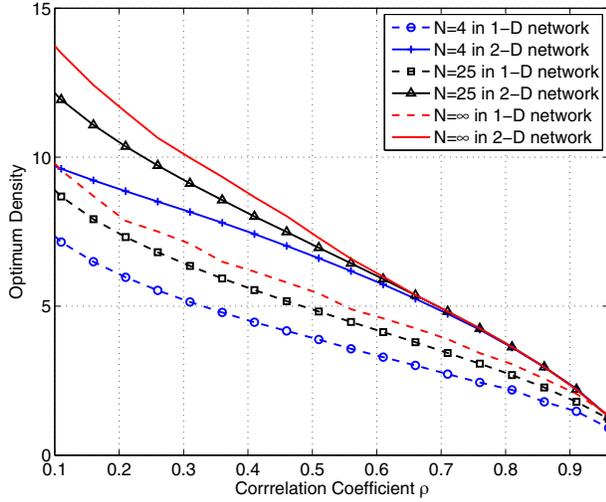


Fig. 8. The optimum or asymptotically optimum node density v.s. spatial correlation coefficient  $\rho$  in the 1-D and 2-D networks ( $\gamma_0 = 10$  dB).

### B. MMSE Spatial Interpolation

The performance of the spatial interpolation in a 2-D network is studied in this subsection. Similar to the 1-D case, we consider the worst case scenario by interpolating the data located in the middle of the square formed by four adjacent sensor nodes, with coordinates of the data points to be estimated being  $\boldsymbol{\eta}'_{ij} = [(i + \frac{1}{2})d, (j + \frac{1}{2})d]^T$ , for  $i, j = 0, \dots, K-1$ . Correspondingly, the data vector can be expressed as  $\boldsymbol{\xi}_d = [\mathbf{x}'_0, \dots, \mathbf{x}'_{K-1}]^T \in \mathcal{R}^{N \times 1}$ , where  $\mathbf{x}'_m = [x(\boldsymbol{\eta}'_{m0}), \dots, x(\boldsymbol{\eta}'_{m(K-1)})]^T \in \mathcal{R}^{K \times 1}$ .

Following the same procedure as in the 1-D case, the error correlation matrix,  $\Phi_{ee}^{(d)} = \mathbb{E}[(\hat{\boldsymbol{\xi}}_d - \boldsymbol{\xi}_d)(\hat{\boldsymbol{\xi}}_d - \boldsymbol{\xi}_d)^T]$ , with  $\hat{\boldsymbol{\xi}}_d$  being the MMSE estimate of  $\boldsymbol{\xi}_d$ , can be calculated by

$$\Phi_{ee}^{(d)} = \Phi_{ss} - \Phi_{ds} \left( \Phi_{ss} + \frac{\delta}{\gamma_0} \mathbf{I}_N \right)^{-1} \Phi_{sd}, \quad (26)$$

where  $\Phi_{dd} = \mathbb{E}[\boldsymbol{\xi}_d \boldsymbol{\xi}_d^T] = \Phi_{ss}$  is used in the above equation,  $\Phi_{ds} = \mathbb{E}[\boldsymbol{\xi}_d \boldsymbol{\xi}_s^T]$ , and  $\Phi_{sd} = \Phi_{ds}^T$ . The cross-correlation matrix,  $\Phi_{ds}$ , can be expressed as

$$\Phi_{ds} = \begin{pmatrix} \mathbf{R}'_0 & \mathbf{R}'_0 & \mathbf{R}'_1 & \cdots & \mathbf{R}'_{(K-2)} \\ \mathbf{R}'_1 & \mathbf{R}'_0 & \mathbf{R}'_0 & \cdots & \mathbf{R}'_{(K-3)} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \mathbf{R}'_{(K-2)} & \mathbf{R}'_{(K-3)} & \cdots & \mathbf{R}'_0 & \mathbf{R}'_0 \\ \mathbf{R}'_{(K-1)} & \mathbf{R}'_{(K-2)} & \cdots & \mathbf{R}'_1 & \mathbf{R}'_0 \end{pmatrix}, \quad (27)$$

where  $\mathbf{R}'_m = \mathbb{E}(\mathbf{x}'_m \mathbf{x}'_0) \in \mathcal{R}^{K \times K}$  is a non-symmetric Toeplitz matrix with the first row being  $[r'_{m,0}, r'_{m,0}, r'_{m,1}, \dots, r'_{m,K-2}]^T \in \mathcal{R}^{k \times 1}$  and the first column  $[r'_{m,0}, r'_{m,1}, \dots, r'_{m,K-1}]^T \in \mathcal{R}^{k \times 1}$ , and  $r'_{m,k} = \mathbb{E}[x(\boldsymbol{\eta}'_{mk})x(\boldsymbol{\eta}_{00})]$  can be expressed as

$$r'_{m,k} = \rho \|\boldsymbol{\eta}'_{mk} - \boldsymbol{\eta}_{00}\| = \rho \sqrt{\frac{(k+\frac{1}{2})^2 + (m+\frac{1}{2})^2}{\delta}}. \quad (28)$$

The matrix  $\Phi_{ds}$  is in the form of a non-symmetric TBT matrix.

From (26), the MSE of the 2-D interpolation in a network with  $N$  nodes can be calculated as  $\vartheta_{d,N}^2 = \frac{1}{N} \text{trace}(\Phi_{ee}^{(d)})$ .

*Proposition 4:* When  $N \rightarrow \infty$  while keeping a finite  $\delta$ , the asymptotic MSE of the data estimated through spatial interpolations in a 2-D network with the spatial correlation coefficient  $\rho$  is

$$\vartheta_d^2 \triangleq \lim_{N \rightarrow \infty} \vartheta_{d,N}^2 = \int_{-\frac{1}{2}}^{\frac{1}{2}} \int_{-\frac{1}{2}}^{\frac{1}{2}} \left[ \Lambda'_{ss}(f_1, f_2) - \frac{|\Lambda'_{ds}(f_1, f_2)|^2}{\Lambda'_{ss}(f_1, f_2) + \frac{\delta}{\gamma_0}} \right] df_1 df_2, \quad (29)$$

where  $\Lambda'_{ss}(f_1, f_2)$  is given in (25), and  $\Lambda'_{ds}(f_1, f_2)$  is computed as

$$\Lambda'_{ds}(f_1, f_2) = \sum_{k=-\infty}^{+\infty} \sum_{m=-\infty}^{+\infty} \rho \sqrt{\frac{(k+\frac{1}{2})^2 + (m+\frac{1}{2})^2}{\delta}} e^{-j2\pi(kf_1 + mf_2)}. \quad (30)$$

*Proof:* The proof is in Appendix F.  $\blacksquare$

Fig. 7 shows the numerical and simulated MSE of the spatial interpolation in the 2-D network, and the results for the 1-D case are also shown in the figure for comparison. The correlation coefficient is  $\rho = 0.8$ . The SNR is  $\gamma_0 = 10$  dB. Simulation results with  $N = 1,600$  are used to approximate the asymptotic results with  $N \rightarrow \infty$ . For both the 1-D and 2-D cases, the MSE is convex in  $\delta$  when  $N$  is small, and it becomes a monotonic decreasing function of  $\delta$  when  $N$  is large. It can be seen that the 2-D network has a worse MSE compared to its 1-D counterpart. The performance difference between the 1-D and 2-D networks becomes smaller as either  $N$  or  $\delta$  increases.

When  $N$  is small such that  $\vartheta_{d,N}^2$  is convex in  $\delta$ , the optimum node density in a 2-D network can be numerically identified by solving  $|\frac{\partial \vartheta_{d,N}^2}{\partial \delta}| = 0$ . When  $N$  is large or tends to infinity, the optimum or asymptotically optimum node density can be obtained by numerically solving  $|\frac{\partial \vartheta_{d,N}^2}{\partial \delta}| = \epsilon$ , with  $\epsilon$  being a very small number. Fig. 8 shows the optimum node density as a function of  $\rho$  in the 1-D and 2-D networks. In the figure,  $\epsilon = 10^{-3}$  is used for both 1-D and 2-D networks when  $N$  is large. For a fixed  $\rho$ , the optimum node density increases as  $N$  increases, and it is upper bounded by the asymptotic result. Therefore, the 1-D and 2-D networks have similar performance trends. For a given  $\rho$ , the optimum node density of the 2-D network is slightly higher than its 1-D counterpart.

## V. CONCLUSIONS

In this paper, the optimum sensor node densities for 1-D and 2-D WSNs with spatial source correlation were studied. The impacts of the node density on the MSE of the data reconstructed at the FC were investigated for both small networks with finite number of nodes, and large networks with infinite number of nodes. Exact analytical expressions of the MSE, many in closed-forms, were obtained for the 1-D and 2-D networks. The analytical results quantitatively identified the interactions among the various system parameters and the estimation fidelity, and the results provide insights and guidelines on the design of practical WSNs.

There were three observations. 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, 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. Finally, the 1-D and 2-D networks have similar performance trends with respect to node density, and their performance difference diminishes as the spatial correlation coefficient increases.

## APPENDIX

### A. Proof of Lemma 1

In Step 2, the MMSE vector  $\mathbf{w}_{sl}$  that minimizes  $\sigma_{\eta}^2$  can be obtained through the orthogonal principal [18],  $\mathbb{E}\{\{\mathbf{w}_{sl}^T \hat{\mathbf{x}}_s - x(\boldsymbol{\eta})\} \hat{\mathbf{x}}_s^T\} = 0$ , and the result is

$$\mathbf{w}_{sl}^T = \sqrt{P_n} \mathbf{r}_{\boldsymbol{\eta}} (P_n \mathbf{R}_{xx} + \sigma_z^2 \mathbf{I}_N)^{-1} \mathbf{W}_s^{-1},$$

where (2) and (6) are used in the above equation. Combining the above equation with (6) and (7) leads to (4).

### B. Proof of Proposition 1

Based on Szego's Theorem [16], when  $N \rightarrow \infty$ , (10) can be rewritten as

$$\sigma_s^2 = \lim_{N \rightarrow \infty} \sigma_{s,N}^2 = \int_{-\frac{1}{2}}^{\frac{1}{2}} \left[ \frac{1}{\Lambda_{ss}(f)} + \frac{\gamma_0}{\delta} \right]^{-1} df, \quad (31)$$

where  $\Lambda_{ss}(f) = \sum_{n=-\infty}^{+\infty} \rho^{|n|d} e^{-jn2\pi f}$  is the discrete-time Fourier transform (DTFT) of the sequence,  $\{\rho^{|n|d}\}_n$ , and it can be calculated by

$$\Lambda_{ss}(f) = \frac{1 - \rho^{2d}}{1 + \rho^{2d} - 2\rho^d \cos(2\pi f)}. \quad (32)$$

Substituting (32) into (31), and solving the integral with [19, eqn. (2.553.3)], we have the result in (11).

### C. Proof of Corollary 1

From (11), 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

$$g_1'(d) = \frac{2\gamma_0}{(1 - \rho^{2d})^2} \times g_2(d, \gamma_0), \quad (33)$$

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

$$g_2(d, \gamma_0) \triangleq (1 - \rho^{2d})^2 (1 + \gamma_0 d) + 2\rho^{2d} (1 - \rho^{2d}) + 4d \log(\rho) \rho^{2d} \quad (34)$$

From (33), 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

$$g_3(v) \triangleq g_2(d, 0) = 1 - v^2 + 2v \log(v), \quad 0 \leq v \leq 1 \quad (35)$$

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 (35), 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.

### D. Proof of Proposition 2

The Toeplitz matrix,  $\mathbf{R}_{ds}$ , is uniquely determined by the sequence  $\mathbf{t}_{ds} = [t_{-(N-1)}, \dots, t_0, \dots, t_{N-1}]^T$ , where  $t_n = \rho^{\frac{d}{2}} \rho^{|n+1|d}$  when  $n < 0$ , and  $t_n = \rho^{\frac{d}{2}} \rho^{nd}$  otherwise. When  $N \rightarrow \infty$ , the DTFT of the sequence  $\mathbf{t}_{ds}$  can be calculated as

$$\Lambda_{ds}(f) = \rho^{\frac{d}{2}} \frac{(1 - \rho^d)(1 + e^{j2\pi f})}{1 + \rho^{2d} - 2\rho^d \cos(2\pi f)}. \quad (36)$$

Based on [16, Lemma 2],  $\mathbf{R}_{ds}$  is asymptotically equivalent to a circulant matrix,  $\mathbf{C}_{ds} = \mathbf{U}_N^H \mathbf{D}_{ds} \mathbf{U}_N$ , where  $\mathbf{U}_N^H$  is the unitary discrete Fourier transform (DFT) matrix with the  $(m, n)$ -th element being  $(\mathbf{D}_{ds})_{m,n} = \frac{1}{\sqrt{N}} \exp[-j2\pi \frac{(m-1)(n-1)}{N}]$ , and  $\mathbf{D}_{ds}$  is a diagonal matrix with its  $k$ -th diagonal element being  $(\mathbf{D}_{ds})_{k,k} = \Lambda_{ds}(\frac{k-1}{N})$ .

Similarly, the Toeplitz matrix,  $\mathbf{R}_{ss}$ , is asymptotically equivalent to a circulant matrix,  $\mathbf{C}_{ss} = \mathbf{U}_N^H \mathbf{D}_{ss} \mathbf{U}_N$ , where  $\mathbf{D}_{ss}$  is a diagonal matrix with its  $k$ -th diagonal element being  $(\mathbf{D}_{ss})_{k,k} = \Lambda_{ss}(\frac{k-1}{N})$ , with  $\Lambda_{ss}(f)$  defined in (32).

Based on [20, Theorem 2.1], the error correlation matrix,  $\mathbf{R}_{ee}^{(d)}$ , is asymptotically equivalent to a circulant matrix,  $\mathbf{C}_{ee}^{(d)} = \mathbf{C}_{ss} - \mathbf{C}_{ds} (\mathbf{C}_{ss} + \frac{\delta}{\gamma_0} \mathbf{I})^{-1} \mathbf{C}_{ds}^H = \mathbf{U}_N^H \mathbf{D}_{ee}^{(d)} \mathbf{U}_N$ , where  $\mathbf{D}_{ee}^{(d)} = \mathbf{D}_{ss} - \mathbf{D}_{ds} (\mathbf{D}_{ss} + \frac{\delta}{\gamma_0} \mathbf{I})^{-1} \mathbf{D}_{ds}^H$ .

Based on Szego's Theorem, we have

$$\sigma_d^2 = \int_{-\frac{1}{2}}^{\frac{1}{2}} \left[ \Lambda_{ss}(f) - \frac{|\Lambda_{ds}(f)|^2}{\Lambda_{ss}(f) + \frac{\delta}{\gamma_0}} \right] df. \quad (37)$$

Substituting (32) and (36) into the above equation and simplifying leads to (17).

### E. Proof of Corollary 3

The MSE in (17) can be alternatively represented as

$$\sigma_d^2 = \{1 + [f_3^{-1}(\delta) - f_3(\delta)] / [f_1^{-1}(\delta) + f_3(\delta)]\}^{\frac{1}{2}}. \quad (38)$$

Since  $f_1(\delta)$  is a decreasing function of  $\delta$  and  $f_3(\delta)$  is an increasing function of  $\delta$ , it is straightforward to show that  $[f_3^{-1}(\delta) - f_3(\delta)] / [f_1^{-1}(\delta) + f_3(\delta)]$  is a decreasing function of  $\delta$ , and this completes the proof.

### F. Proof of Proposition 4

According to [17, Lemma 1], the TBT matrices,  $\Phi_{ss}$  and  $\Phi_{ds}$ , are asymptotically equivalent to circulant-block-circulant (CBC) matrices,  $\mathbf{B}_{ss}$  and  $\mathbf{B}_{ds}$ , respectively, where the eigenvalues of  $\mathbf{B}_{ss}$  and  $\mathbf{B}_{ds}$  are samples of  $\Lambda'_{ss}(f_1, f_2)$  and  $\Lambda'_{ds}(f_1, f_2)$ , respectively [17, Theorem 3]. In addition, the CBC matrices,  $\mathbf{B}_{ss}$  and  $\mathbf{B}_{ds}$ , share the same orthonormal eigenvectors [21]. Once the asymptotic equivalence is established, the rest of the proof follows the same procedure as described in Appendix D for the 1-D case.

## REFERENCES

- [1] Q. Zhao and L. Tong, "Energy-efficient information retrieval for correlated source reconstruction in sensor networks," *IEEE Trans. Wireless Commun.*, vol. 6, pp. 157-165, Jan. 2007.
- [2] A. Anandkumar, L. Tong, and A. Swami, "Optimal node density for detection in energy-constrained random networks," *IEEE Trans. Signal Process.*, vol. 56, pp. 5232-5245, Oct. 2008.

- [3] T. Wimalajeewa and S. K. Jayaweera, "Impact of mobile node density on detection performance measures in a hybrid sensor network," *IEEE Trans. Wireless Commun.*, vol. 9, pp. 1760–1769, May 2010.
- [4] W. Wang, V. Srinivasan, B. Wang, and K.-C. Chua, "Coverage for target localization in wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 667–676, Feb. 2008.
- [5] B. Wang, K.-C. Chua, V. Srinivasan, and W. Wang, "Information coverage in randomly deployed wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 6, pp. 2994–3004, Aug. 2007.
- [6] L. Liu, X. Zhang, and H. Ma, "Localization-oriented coverage in wireless camera sensor networks," *IEEE Trans. Wireless Commun.*, vol. 10, pp. 484–494, Feb. 2011.
- [7] J. F. Chamberland and V. V. Veeravalli, "How dense should a sensor network be for detection with correlated observations?" *IEEE Trans. Inf. Theory*, vol. 52, pp. 5099–5106, Nov. 2006.
- [8] R. Cristescu, and M. Vetterli, "On the optimal density for real-time data gathering of spatio-temporal processes in sensor networks," in *Proc. 2005 IPSN*, pp. 159–164.
- [9] X. Zhang, H. Wang, F. N. Abdesselam, and A. A. Khokhar, "Distortion analysis for real-time data collection of spatially temporally correlated data fields in wireless sensor networks," *IEEE Trans. Veh. Technol.*, vol. 58, pp. 1583–1594, Mar. 2009.
- [10] M. Gastpar, and M. Vetterli, "Power, spatio-temporal bandwidth, and distortion in large sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 23, pp. 745–755, Apr. 2005.
- [11] Y. Sung, H. V. Poor, and H. Yu, "Optimal node density for two-dimensional sensor arrays," in *Proc. 2008 IEEE Sensor Array Multi-channel Signal Process. Workshop*, pp. 271–274.
- [12] J. Wu and N. Sun, "Optimal sensor density in a distortion-tolerant linear wireless sensor network," in *Proc. 2010 IEEE Global Telecommun. Conf.*
- [13] P. Gupta and P. R. Kumar, "The capacity of wireless networks," *IEEE Trans. Inf. Theory*, vol. 46, pp. 388–404, Mar. 2000.
- [14] S. D. Servetto, "On the feasibility of large scale wireless sensor networks," in *Proc. 2002 IEEE Int. Conf. Commun., Control, Comput.*
- [15] 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," in *Proc. 2003 Intern. Conf. Inf. Process. Sensor Netw.*, pp. 1–16.
- [16] 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. Inf. Theory*, vol. 47, pp. 1243–1251, Mar. 2001.
- [17] P. A. Voois, "A theorem on the asymptotic eigenvalue distribution of Toeplitz-block-Toeplitz matrix," *IEEE Trans. Signal Process.*, vol. 44, pp. 1837–1841, July 1996.
- [18] S. M. Kay, *Fundamentals of Statistical Signal Processing, Vol. I: Estimation Theory*. Prentice-Hall, 1993.
- [19] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 6th edition. Academic Press, 2000.
- [20] R. M. Gray, *Toeplitz and Circulant Matrices: A Review*. NOW Publishers, 2006.
- [21] T. D. Mazancourt and D. Gerlic, "The inverse of a block-circulant matrix," *IEEE Trans. Antennas Propag.*, vol. AP-31, pp. 808–810, Sep. 1983.



**Jingxian Wu** (S'02-M'06) received the B.S. degree in electronic engineering from Beijing University of Aeronautics and Astronautics, Beijing, China, in 1998, the M.S. degree in electronic engineering from Tsinghua University, Beijing, China, in 2001, and the Ph.D. degree in electrical engineering from the University of Missouri, Columbia, in 2005.

He is currently an Assistant Professor with the Department of Electrical Engineering, University of Arkansas, Fayetteville. His research interests mainly focus on wireless communications and wireless networks,

including ultra-low power communications, cooperative communications, cognitive radio, and cross-layer optimization, etc. He is an Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, and served as an Associate Editor of the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY from 2007 to 2011. He served as a Cochair for the 2012 Wireless Communication Symposium of the IEEE International Conference on Communication, and a Cochair for the 2009 Wireless Communication Symposium of the IEEE Global Telecommunications Conference. Since 2006, he has served as a Technical Program Committee Member for a number of international conferences, including the IEEE Global Telecommunications Conference, the IEEE Wireless Communications and Networking Conference, the IEEE Vehicular Technology Conference, and the IEEE International Conference on Communications.



**Ning Sun** received the B.S. and M.S. degrees in Electrical Engineering from Shandong University of Science and Technology, Qingdao, China, in 2005 and 2008, respectively. From 2008 to 2009, she was a research assistant at the Underwater Acoustic Communication Institute, Soongsil University, Seoul, South Korea. She is currently a PhD candidate in the Department of Electrical Engineering, University of Arkansas, Fayetteville, USA. She is a student member of IEEE. Her research interests include wireless communications, wireless sensor

networks and digital image processing.