Scaling Laws for Dense Gaussian Sensor Networks and the Order Optimality of Separation

Nan Liu, *Student Member, IEEE*, and Sennur Ulukus, *Member, IEEE*

***Abstract*—We investigate the optimal performance of dense sensor networks by studying the joint source–channel coding problem. There are N uniformly spaced sensor nodes sampling noiselessly a one-dimensional spatial random process over an interval $[0, U_0]$. The overall goal of the sensor network is for the sensor nodes to code and transmit the measurement samples to a collector node over a cooperative multiple-access channel with noisy feedback, and for the collector node to reconstruct the entire random process with minimum expected distortion. We provide separation-based lower and upper bounds for the minimum achievable expected distortion when the underlying random process is Gaussian. When the Gaussian random process satisfies some general conditions, such as the eigenvalues of its Karhunen–Loeve expansion decrease roughly inverse polynomially in order x, i.e., the kth eigenvalue is roughly k^{-x}, we evaluate the lower and upper bounds explicitly, and show that they are of the same order for a wide range of power constraints. Thus, for these random processes, under these power constraints, we show that the minimum achievable expected distortion decreases as $(\log NP(N))^{1-x}$, where $P(N)$ is the sum power constraint on the sensor nodes. Further, we show that the achievability scheme that achieves the lower bound on the distortion is a separation-based scheme that is composed of multiterminal rate-distortion coding and amplify-and-forward channel coding. Therefore, we conclude that separation is order-optimal for the dense Gaussian sensor network scenario under consideration, when the underlying random process satisfies some general conditions.**

***Index Terms*—Cooperation, correlation, scaling laws, sensor networks, separation principle.**

I. Introduction

With the recent advances in the hardware technology, small cheap nodes with sensing, computing and communication capabilities have become available. In practical applications, it is possible to deploy a large number of these nodes to sense the environment. In this paper, we investigate the optimal performance of a dense sensor network by studying the joint source–channel coding problem. The sensor network is composed of N sensors, where N is very large, and a single collector node. Each sensor node has the capability of taking noiseless samples from an underlying random process. Each node in the sensor network is equipped with one transmit and one receive antenna to transmit and receive signals through the wireless medium, i.e., all nodes hear a linear combination of the signals transmitted by all other nodes at that time instant. The overall goal of the sensor network is to take measurements from an underlying random process $S(u)$, $0 \leq u \leq U_0$, code and transmit those measured samples to a collector node, which wishes to reconstruct the entire random process with as little distortion as possible; see Fig. 1. Due to the small distances between the sensor nodes and the correlation in the measured data, the underlying sensor samples are correlated, and due to the existence of receive antennas at the sensor nodes and a transmit antenna at the collector node, the communication channel is a Gaussian cooperative multiple-access channel with noisy feedback. We investigate the minimum achievable expected distortion and a corresponding achievability scheme when the underlying random process is Gaussian.

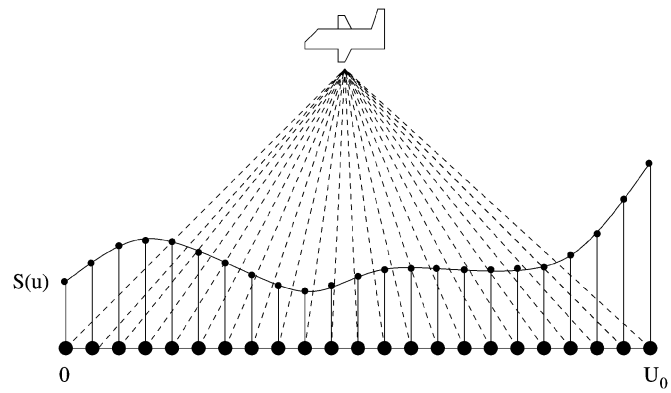

Fig. 1. Sensor network.

Following the seminal paper of Gupta and Kumar [3], which showed that multihop wireless *ad hoc* networks, where users transmit independent data and utilize single-user coding, decoding, and forwarding techniques, do not scale successfully, Scaglione and Servetto [4] investigated the scalability of the sensor networks. Sensor networks, where the observed data is correlated, may scale successfully for two reasons: first, the correlation among the sampled data increases with the increasing number of nodes and, hence, the amount of information the network needs to carry does not increase as fast as in *ad hoc* wireless networks; and second, correlated data facilitates cooperation, and may increase the information-carrying capacity of the network. The goal of the sensor network in [4] was that each sensor reconstructs the data measured by all of the sensors using sensor broadcasting. In this paper, we focus on the case where

Manuscript received September 1, 2006; revised January 14, 2007. This work was supported by the National Science Foundation under Grants CCR 03-11311, CCF 04-47613, and CCF 05-14846. The material in this paper was presented in part at the IEEE International Conference on Communications, Istanbul, Turkey, June 2006 and at the IEEE International Symposium on Information Theory, Seattle, WA, July 2006.

The authors are with the Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20742 USA (e-mail: nkancy@umd.edu; ulukus@umd.edu).

Communicated by J. N. Laneman, Guest Editor for the Special Issue on Relaying and Cooperation.

Digital Object Identifier 10.1109/TIT.2007.904966

0018-9448/$25.00 © 2007 IEEE

the reconstruction is required only at the collector node. Also, in this paper, the task is not the reconstruction of the data the sensors measured, but the reconstruction of the underlying random process.

Marco *et al.* [5] is the first paper to formulate the sensor network problem considered here, where there is a single collector node which wishes to reconstruct the random process; see also [6]. The channel model used in [5] was similar to that used in [3], and is interference limited. The sensor encoders were limited to scalar quantization with entropy-rate coding. It was shown that the system performance becomes asymptotically poor as the number of sensors grows, i.e., the sensor network under consideration does not scale successfully.

El Gamal [7] studied the same problem as in [5], but removed the constraint that the channel model is interference limited. By modeling the channel as a cooperative Gaussian multiple-access channel, [7] showed that all spatially band-limited Gaussian processes can be estimated at the collector node, subject to any nonzero constraint on the mean squared distortion, i.e., the sensor network scales successfully. In this paper, we study the minimum achievable expected distortion for space-limited, and thus, not band-limited, random processes, and we show the rate at which the minimum achievable expected distortion decreases to zero as the number of nodes increases. Also, in [7], it is assumed that the channel gains between the nodes are decreasing functions of the distance between them, without enforcing any upper bounds. This implies that, when the sensors are placed very densely, the channel gains between nearby sensors become unboundedly large. This physically impossible situation arises because although the channel model used in [7] is valid only in the far field of the transmitter, it is used for all distances. Although we adopted this channel model in [1], an earlier conference version of this work where the underlying random process was Gauss–Markov, we have changed the channel model to a more realistic one in this paper (and [2], the conference version of this work), where we assume that the channel gains decrease with distance, however, they are lower and upper bounded. The difference in the channel models in [1] and here (and [2]), does not affect our conclusion, i.e., in both cases, we are able to find achievable schemes that achieve the lower bound on the distortion in the order sense. However, it affects the achievability scheme itself; in [1], the achievability scheme is based on the basic idea of decode-and-forward as the channel model allows a significant number of nodes to be able to decode successfully the signal transmitted by a node, whereas the achievability scheme here (and [2]) is based on the basic idea of amplify-and-forward, where due to the lower and upper bounds on the channel gains, a sufficient amount of beamforming effect is achieved through the amplify-and-forward scheme.

Kashyap *et al.* [8] studied the source coding part of the problem investigated in this paper. The paper showed that for any distortion constraint that is independent of N, the difference between the rate achievable by distributed source coding and the rate achievable by centralized source coding is bounded by a constant, independent of N. Though we study a joint source–channel coding problem, both our converse and achievability proofs are separation-based, and thus, we show a similar result: in the source coding part we show that the ratio between the rate achievable by distributed source coding and the rate achievable by centralized source coding is bounded by a constant, independent of N. In contrast to [8], where the distortion constraint is independent of N, we allow the distortion to go to zero as a function of N. Moreover, [8] deals with stationary Gaussian random processes, while we allow for nonstationarity of the underlying random process. It is not immediately evident whether the methods in [8] apply in the scenario considered in this paper.

Neuhoff and Pradhan [9] studied the source-coding part of the problem investigated in this paper by allowing the random process to be unbounded in space. The sensors are densely as well as widely distributed. In this case, results from Grenander and Szego [10] were used. However, for the case of a finite interval, as considered in this paper, such results cannot be used.

Gastpar and Vetterli [11] studied the case where the sensors observe a noisy version of a linear combination of L Gaussian random variables which all have the same variance, code, and transmit those observations to a collector node, and the collector node reconstructs the L random variables. In [11], the expected distortion achieved by applying separation-based approaches was shown to be order worse than the lower bound on the minimum achievable expected distortion. In this paper, we study the case where the data of interest at the collector node is not a finite number of random variables, but a random process, which, using Karhunen–Loeve expansion, can be shown to be equivalent to a set of infinitely many random variables with varying variances. We assume that the sensors are able to take noiseless samples, but that each sensor observes only its own sample. Our upper bound on the minimum achievable distortion is also developed by using a separation-based approach, but it is shown to be of the same order as the lower bound, for a wide range of power constraints, for random processes that satisfy some general conditions.

From an information-theoretic point of view, our problem is a joint source–channel coding problem for lossy communication of correlated sources over a cooperative Gaussian multiple-access channel with noisy feedback. The simpler problem of lossless reconstruction of correlated sources over a multiple-access channel without cooperation or feedback still remains open [12]–[15]. Therefore, a direct and closed-form expression for the distortion seems unlikely to be obtained, and, consequently, we resort to developing lower and upper bounds. We first provide lower and upper bounds for the minimum achievable expected distortion for arbitrary Gaussian random processes whose Karhunen–Loeve expansion exists. Then, we focus on the case where the Gaussian random process also satisfies some general conditions, such as the eigenvalues of its Karhunen–Loeve expansion decreases roughly inverse polynomially in order x, i.e., the kth eigenvalue is roughly k^{-x}. For these random processes, we evaluate the lower and upper bounds explicitly, and show that they are of the same order, for a wide range of power constraints. Thus, for these random processes, under a wide range of power constraints, we determine an order-optimal achievability scheme, and identify the minimum achievable expected distortion as a function of the number of nodes and the sum power constraint. We show

that the minimum achievable expected distortion decreases to zero at the rate of $(\log NP(N))^{1-x}$, where $P(N)$ is the sum power constraint on the sensor nodes. Our achievability scheme is separation-based: each sensor node first performs multiterminal source coding [16], then, performs channel coding, and utilizes the cooperative nature of the wireless medium through the amplify-and-forward scheme [17]. In multiuser information theory, generally speaking, the separation principle does not hold. However, in our case, we have found a scheme which is separation based, and is order-optimal.

II. System Model

The collector node wishes to reconstruct a random process $S(u)$, for $0 \leq u \leq U_0$, where u denotes the spatial position; $S(u)$ is assumed to be Gaussian with zero-mean and a continuous autocorrelation function $K(u,v)$. The N sensor nodes are placed on a straight line at positions $0 \leq u_1 \leq u_2 \leq \cdots \leq u_N \leq U_0$, and observe samples

$$\boldsymbol{S}_N = [\, S(u_1) \quad S(u_2) \quad S(u_N)]^T. \tag{1}$$

For simplicity and to avoid irregular cases, we assume that the sensors are equally spaced, i.e.,

$$u_i = \frac{i-1}{N} U_0, \qquad i = 1, 2, \ldots, N. \tag{2}$$

The distortion measure is the squared error

$$d\left(S(u), \hat{S}(u)\right) = \frac{1}{U_0} \int_0^{U_0} \left(S(u) - \hat{S}(u)\right)^2 du. \tag{3}$$

Each sensor node and the collector node, denoted as node 0, is equipped with one transmit and one receive antenna. To simplify the presentation, from now until Section VII, we will assume that the collector node does not use its transmit antenna, and thus, there is no feedback in the system. We will allow the collector node to use its transmit antenna and provide feedback to the sensor nodes in Section VII, and show that the results of the previous sections remain unchanged. The transmissions through the wireless medium are time slotted. The channel is assumed to be memoryless between the time slots. At any time instant, let X_i denote the signal transmitted by node i, and Y_j denote the signal received at node j. Let h_{ij} denote the channel gain from node i to node j. Then, the received signal at node j can be written as

$$Y_j = \sum_{i=1, i \neq j}^{N} h_{ij} X_i + Z_j, \qquad j = 0, 1, 2, \ldots, N \tag{4}$$

where $\{Z_j\}_{j=0}^N$ is a vector of $N+1$ independent and identically distributed (i.i.d.), zero-mean, unit-variance Gaussian random variables. Therefore, the channel model of the network is such that all nodes hear a linear combination of the signals transmitted by all other nodes at that time instant. We assume that the channel gain h_{ij} is bounded, i.e.,

$$\bar{h}_l \leq h_{ij} \leq \bar{h}_u, \qquad i = 1, \ldots, N, \;\; j = 0, 1, \ldots, N \tag{5}$$

where $\bar{h}_u$ and $\bar{h}_l$ are positive constants independent of N. This model is very general and should be satisfied very easily. By the conservation of energy, $h_{ij}^2 \leq 1$, and since all nodes are within finite distances of each other, the channel gains should be lower bounded as well.

We assume that all sensors share the sum power constraint $P(N)$. The two most interesting cases for $P(N)$ are $P(N) = NP_{\text{ind}}$, where each sensor has its individual power constraint P_{ind}, and $P(N) = P_{\text{tot}}$, where the sum power constraint is a constant P_{tot} and does not depend on the number of sensor nodes. In the latter case, when more and more sensor nodes are deployed, the individual power of each sensor node decreases as P_{tot}/N. Our goal is to determine the scheme that achieves the minimum achievable expected distortion

$$\frac{1}{U_0} \int_0^{U_0} E\left[\left(S(u) - \hat{S}(u)\right)^2\right] du \tag{6}$$

at the collector node for a given sum power constraint $P(N)$, and also to determine the rate at which this distortion goes to zero as a function of the number of sensor nodes and the sum power constraint.

Next, we give a more precise definition of our problem. Each sensor node observes a sample of a sequence of spatial random processes $\{S^{(l)}(u)\}_{l=1}^n$ i.i.d. in time, where index l denotes time, u denotes the spatial position, and n is the block length of the sequence of random processes, and also the delay parameter, which may be a function of N, the number of sensor nodes. For now, we assume that n channel uses are allowed for n realizations of the random process; the case where we allow the number of channel uses and the number of realizations to differ will be treated in Section VII. At time instant m, sensor node j transmits

$$X_j(m) = F_j^{(m)}\left(\left\{S^{(l)}(u_j)\right\}_{l=1}^n, \left\{Y_j^{(l)}\right\}_{l=n+1}^{m-1}\right), \qquad m = n+1, n+2, \ldots, 2n, \;\; j = 1, 2, \ldots, N \tag{7}$$

i.e., after waiting n time slots to obtain a block of observations, the sensor node transmits, at time m, a signal that is a function of its observations of the entire block of random process samples and also the signal it received before time m. We are interested in the performance in the information-theoretic sense and, hence, we allow the delay n to be arbitrarily large. By the assumption of sum power constraint, we have

$$\frac{1}{n} \sum_{m=1}^{n} \sum_{j=1}^{N} E\left[X_j^2(m)\right] \leq P(N). \tag{8}$$

The collector node reconstructs the random process as

$$\left\{\hat{S}^{(l)}(u), u \in [0, U_0]\right\}_{l=1}^n = G\left(Y_0^{(n+1)}, Y_0^{(n+2)}, \ldots, Y_0^{(2n)}\right). \tag{9}$$

For fixed encoding functions of the nodes $\left\{F_j^{(m)}\right\}_{m=n+1, j=1}^{m=2n, j=N}$ and the decoding function of the collector node G, the achieved expected distortion is

$$D^N = \frac{1}{n} \sum_{l=1}^{n} E\left[d\left(S^{(l)}(u), \hat{S}^{(l)}(u)\right)\right] \tag{10}$$

and we are interested in the smallest achievable expected distortion over all encoding and decoding functions where n is allowed to be arbitrarily large.

In this paper, our purpose is to understand the behavior of the minimum achievable expected distortion when the number of sensor nodes is very large. We introduce the big-O, big-Ω, and big-Θ notations. We say that f is $O(g)$, and g is $\Omega(f)$, if there exist constants c and k, such that $|f(N)| \leq c|g(N)|$ for all $N > k$; we say that f is $\Theta(g)$ if $f(N)$ is both $O(g)$ and $\Omega(g)$. All logarithms are defined with respect to base e, and $\lfloor x \rfloor$ denotes the largest integer smaller than or equal to x. $(\boldsymbol{x})_i$ and $\|\boldsymbol{x}\|$ denote the ith element and the Euclidean norm of vector $\boldsymbol{x}$, respectively. $\|\boldsymbol{A}\|_2$ denotes the spectral norm of matrix $\boldsymbol{A}$, which is defined as the square root of the largest eigenvalue of matrix $\boldsymbol{A}^T\boldsymbol{A}$ [18].

III. A Class of Gaussian Random Processes

For a Gaussian random process $S(u)$ with a continuous autocorrelation function, we perform the Karhunen–Loeve expansion [19]

$$S(u) = \sum_{k=0}^{\infty} \bar{S}_k \phi_k(u) \tag{11}$$

to obtain the ordered eigenvalues $\{\lambda_k\}_{k=0}^{\infty}$, and the corresponding set of orthonormal eigenfunctions $\{\phi_k(u), u \in [0, U_0]\}_{k=0}^{\infty}$.

Let $\mathcal{A}$ be the set of Gaussian random processes on $[0, U_0]$ with continuous autocorrelation functions, that satisfy the following conditions.

1. There exist nonnegative constants d_l, d_u, and nonnegative integers $c_l, c_u, K_0 \geq c_u+1$, and two sequences of numbers $\{\lambda'_k\}_{k=0}^{\infty}$ and $\{\lambda''_k\}_{k=0}^{\infty}$ defined as

$$\lambda'_k = \begin{cases} \lambda_k, & k \leq K_0 \\ \frac{d_l}{(k+c_l)^x}, & k > K_0 \end{cases} \tag{12}$$

and

$$\lambda''_k = \begin{cases} \lambda_k, & k \leq K_0 \\ \frac{d_u}{(k-c_u)^x}, & k > K_0 \end{cases} \tag{13}$$

for some constant $x > 1$, such that

$$\lambda'_k \leq \lambda_k \leq \lambda''_k. \tag{14}$$

The condition that $x > 1$ is without loss of generality, because for all continuous autocorrelations, the eigenvalues decrease faster than k^{-1}.

2. In addition to continuity, $K(u, v)$ satisfies the Lipschitz condition of order $1/2 < \alpha \leq 1$, i.e., there exists a constant $B > 0$ such that

$$|K(u_1, v_1) - K(u_2, v_2)| \leq B\left(\sqrt{(u_1-u_2)^2 + (v_1-v_2)^2}\right)^{\alpha} \tag{15}$$

for all $u_1, v_1, u_2, v_2 \in [0, U_0]$.

3. For $k = 0, 1, \ldots$, the function $\phi_k(v)$ and the function $K(u, v)\phi_k(v)$ as a function of v satisfy the following condition: there exist positive constants B_1, B_2, B_3, B_4, $\beta \leq 1$, $\gamma \leq 1$, and nonnegative constant τ, independent of k, such that

$$|\phi_k(v_1) - \phi_k(v_2)| \leq B_3(k + B_4)^{\tau}|v_1 - v_2|^{\gamma} \tag{16}$$

and

$$|K(u, v_1)\phi_k(v_1) - K(u, v_2)\phi_k(v_2)| \leq B_2(k + B_1)^{\tau}|v_1 - v_2|^{\beta} \tag{17}$$

for all $u, v_1, v_2 \in [0, U_0]$.

The reasons why these conditions are needed for the explicit evaluation of the lower and upper bounds on the minimum achievable expected distortion will be clear from the proofs. Here, we provide some intuition as to why they are needed. Condition 1 states that we consider random processes that have eigenvalues λ_k which decrease at a rate of approximately k^{-x}. The rate of decrease in the eigenvalues is an indication of how the randomness of the random process is distributed upon the eigenfunctions. For example, a small x means that the randomness is distributed rather evenly upon all eigenfunctions, while a large x means that the randomness is mostly concentrated upon a subset of eigenfunctions. Thus, the minimum achievable expected distortion depends crucially on the rate of decrease parameter x. The lower (upper) bound on the eigenvalues in (14) will be used to calculate the lower (upper) bound on the minimum achievable expected distortion. Conditions 2 and 3 are needed because instead of the random process itself that is of interest to the collector node, the collector node, at best, can know only the sampled values of the random process. How well the entire process can be approximated from its samples is of great importance in obtaining quantitative results. Lipschitz conditions describe the quality of this approximation well. By Condition 3, we require the variation in the eigenfunction ϕ_k to be no faster than k^{τ}. We note that the well-known trigonometric basis satisfies this condition.

We also note that our conditions are quite general. Many random processes satisfy these conditions, including the Gauss–Markov process, Brownian motion process, centered Brownian bridge, etc. For example, a Gauss–Markov process, also known as the Ornstein–Uhlenbeck process [20], [21], is defined as a random process that is stationary, Gaussian, Markovian, and continuous in probability. It is known that the autocorrelation function of this process is [22]–[24]

$$K(u, v) = \frac{\sigma^2}{2\eta} e^{-\eta|u-v|}. \tag{18}$$

The Karhunen–Loeve expansion of the Gauss–Markov process yields the eigenfunctions $\{\phi_k(u)\}_{k=0}^{\infty}$

$$\phi_k(u) = b_k \left(\cos\sqrt{\frac{\sigma^2}{\lambda_k} - \eta^2}\, u + \frac{\eta}{\sqrt{\frac{\sigma^2}{\lambda_k} - \eta^2}} \sin\sqrt{\frac{\sigma^2}{\lambda_k} - \eta^2}\, u \right) \tag{19}$$

where $\{\lambda_k\}_{k=0}^{\infty}$ are the corresponding eigenvalues and b_k are positive constants chosen such that the eigenfunctions $\phi_k(u)$

have unit energy. It can be shown that $\{\lambda_k\}_{k=0}^{\infty}$ may be bounded as

$$\lambda'_k \leq \lambda_k \leq \lambda''_k \tag{20}$$

where $\{\lambda'_k\}_{k=0}^{\infty}$ is defined as

$$\lambda'_k = \begin{cases} \lambda_k, & k \leq K_0 \\ \frac{\sigma^2 U_0^2}{(k+1)^2\pi^2}, & k > K_0 \end{cases} \tag{21}$$

with $K_0 = \max\left(2, \left\lfloor \frac{\eta^2 U_0^2}{\pi^2} - \frac{3}{4} \right\rfloor\right)$, and $\{\lambda''_k\}_{k=0}^{\infty}$ is defined as

$$\lambda''_k = \begin{cases} \lambda_k, & k \leq K_0 \\ \frac{\sigma^2 U_0^2}{(k-1)^2\pi^2}, & k > K_0. \end{cases} \tag{22}$$

Thus, we observe that the Gauss–Markov process satisfies the conditions defined in this section with $x = 2$, $d_l = d_u$, and $\alpha = \beta = \tau = \gamma = 1$. In fact, in a preliminary conference version of our work [1], we focused specifically on the Gauss–Markov process and presented results similar to those here. We also note, as discussed in the Introduction, that the channel model in [1] is somewhat different than here, and therefore the order-optimal achievability schemes in [1] and here are different.

The lower and upper bounds on the minimum achievable expected distortion will be calculated using $\{\lambda'_k\}_{k=0}^{\infty}$ and $\{\lambda''_k\}_{k=0}^{\infty}$, respectively. Some properties of $\{\lambda'_k\}_{k=0}^{\infty}$ and $\{\lambda''_k\}_{k=0}^{\infty}$ which will be used in later proofs are stated in Lemmas A.1 and A.2 and proved in Appendix I-A.

IV. A Lower Bound on the Achievable Distortion

A. Arbitrary Gaussian Random Processes

A lower bound is obtained by assuming that all of the sensor nodes know the random process exactly, i.e., $S(u), u \in [0, U_0]$, and the sensor network forms an N-transmit 1-receive antenna point-to-point system to transmit the random process to the collector node. Let C_u^N be the capacity of this point-to-point system in nats per channel use and $D_p(R)$ be the distortion–rate function of the random process $S(u)$[25]. In this point-to-point system, the separation principle holds, and therefore

$$D^N \geq D_p\left(C_u^N\right). \tag{23}$$

To evaluate $D_p\left(C_u^N\right)$, we first find the distortion–rate function $D_p(R)$ of $S(u)$ [25, Sec. 4.5] as

$$R(\theta) = \sum_{k=0}^{\infty} \max\left(0, \frac{1}{2}\log\left(\frac{\lambda_k}{\theta}\right)\right) \tag{24}$$

and

$$D(\theta) = U_0^{-1} \sum_{k=0}^{\infty} \min(\theta, \lambda_k) \tag{25}$$

where θ is an intermediate variable used to describe the distortion–rate function. The distortion–rate function $D_p(R)$ characterizes the minimum achievable expected distortion when we use R nats per source realization to describe the random process.

Next, we find C_u^N, the capacity of the N-transmit 1-receive antenna point-to-point system [26] as

$$C_u^N = \frac{1}{2}\log\left(1 + \sum_{i=1}^{N} h_{i0}^2 P(N)\right) \text{ nats/channel use.} \tag{26}$$

To see how C_u^N changes with N, using (26) and (5), we can lower and upper bound C_u^N as

$$\frac{1}{2}\log\left(1 + \bar{h}_l^2 NP(N)\right) \leq C_u^N \leq \frac{1}{2}\log\left(1 + \bar{h}_u^2 NP(N)\right). \tag{27}$$

For arbitrary Gaussian random processes, a lower bound on the minimum achievable expected distortion is

$$D_l^N = D_p\left(C_u^N\right) \tag{28}$$

B. The Class of Gaussian Random Processes in $\mathcal{A}$

Next, we evaluate $D_p\left(C_u^N\right)$ for the class of Gaussian random processes in $\mathcal{A}$. Based on the structure of the eigenvalues in (12) and (14), and the properties of $\{\lambda'_k\}_{k=0}^{\infty}$ in Lemma A.1 in Appendix I-A, the rate–distortion function of the random process satisfies the following lemma.

Lemma 4.1: For Gaussian random processes in $\mathcal{A}$, for any constant $0 < \kappa < 1$, we have

$$R(\theta) \geq \frac{\kappa x d_l^{\frac{1}{x}}}{2}\theta^{-\frac{1}{x}} \tag{29}$$

$$D(\theta) \geq \kappa\left(1 + \frac{\kappa}{x-1}\right)\frac{d_l^{\frac{1}{x}}}{U_0}\theta^{1-\frac{1}{x}} \tag{30}$$

when θ is small enough.

A proof of Lemma 4.1 is provided in Appendix I-B. Using Lemma 4.1, and recognizing the facts that $D(\theta)$ is a nondecreasing function of θ, and $R(\theta)$ is a strictly decreasing function of θ when $\theta < \lambda_1$, i.e., its inverse function $\theta(R)$ exists when R is large enough, we have the next theorem which presents a lower bound for the distortion–rate function of the random process.

Theorem 4.1: For Gaussian random processes in $\mathcal{A}$, for any constant $0 < \kappa < 1$, we have

$$D_p(R) \geq \kappa\left(1 + \frac{\kappa}{x-1}\right)\left(\frac{\kappa x}{2}\right)^{x-1}\frac{d_l}{U_0}R^{1-x} \tag{31}$$

when R is large enough.

We will divide our discussion into two separate cases based on the sum power constraint $P(N)$. For the first case, $P(N)$ is such that

$$\lim_{N\to\infty}\frac{1}{NP(N)} = 0. \tag{32}$$

The cases $P(N) = NP_{\text{ind}}$ and $P(N) = P_{\text{tot}}$ are included in $P(N)$ satisfying (32). From (27), we see that in this case C_u^N increases monotonically in N. Hence, when N is large enough, C_u^N will be large enough such that Theorem 4.1 holds. Hence, for any constant $0 < \kappa < 1$, a lower bound on the minimum achievable expected distortion is

$$\begin{aligned} D_l^N &= D_p\left(C_u^N\right) \\ &\geq \kappa^2\left(1 + \frac{\kappa}{x-1}\right)(\kappa x)^{x-1}\frac{d_l}{U_0}\left(\frac{1}{\log\left(NP(N)\right)}\right)^{x-1} \end{aligned} \tag{33}$$

when N is large enough. Hence, when sum power constraint $P(N)$ satisfies (32), the minimum achievable distortion is

$$\Omega\left(\left(\frac{1}{\log(NP(N))}\right)^{x-1}\right). \tag{34}$$

For the second case, $P(N)$ is such that (32) is not satisfied. In this case, C_u^N is either a constant independent of N or goes to zero as N goes to infinity. The minimum achievable distortion does not go to zero with increasing N.

Therefore, for all possible sum power constraints $P(N)$, the minimum achievable distortion is

$$\Omega\left(\min\left(\left(\frac{1}{\log(NP(N))}\right)^{x-1}, 1\right)\right). \tag{35}$$

When the sum power constraint $P(N)$ grows almost exponentially with the number of nodes, the lower bound on the minimum achievable expected distortion in (35) decreases inverse polynomially with N. Even though this provides excellent distortion performance, it is impractical since sensor nodes are low-energy devices and it is often difficult, if not impossible, to replenish their batteries. When the sum power constraint $P(N)$ is such that (32) is not satisfied, the transmission power is so low that the communication channels between the sensors and the collector node are as if they did not exist. From (35), the lower bound on the estimation error in this case is on the order of 1, which is equivalent to the collector node blindly estimating $S(u) = 0$ for all $u \in [0, U_0]$. Even though the consumed sum power $P(N)$ is very low in this case, the performance of the sensor network is unacceptable; even the lower bound on the minimum achievable expected distortion does not decrease to zero with the increasing number of nodes. For practically meaningful sum power values, including the cases of $P(N) = NP_{\text{ind}}$ and $P(N) = P_{\text{tot}}$, the lower bound on the minimum achievable expected distortion in (35) decays to zero at the rate of

$$\frac{1}{(\log N)^{x-1}}. \tag{36}$$

V. An Upper Bound on the Achievable Distortion

A. Arbitrary Gaussian Random Processes

Any distortion found by using any achievability scheme will serve as an upper bound for the minimum achievable expected distortion. We consider the following separation-based achievable scheme. First, we perform multiterminal rate–distortion coding at all sensor nodes using [16, Theorem 1]. After obtaining the indices of the rate–distortion codes, we transmit the indices as independent messages using the amplify-and-forward method introduced in [17]. The distortion obtained using this scheme will be denoted as D_u^N.

First, we determine an achievable rate region for the communication channel from the sensor nodes to the collector node. The channel by its nature is a multiple-access channel with potential cooperation between the transmitters. The capacity region for this channel is not known. We get an achievable rate region for this channel by using the idea presented in [17].

Theorem 5.1: When the sum power constraint $P(N)$ is such that there exists an $\epsilon > 0$ where

$$\lim_{N\to\infty} P(N)N^{\frac{1}{2}-\epsilon} > 1 \tag{37}$$

for any constant $0 < \kappa < 1$, the following rate region is achievable:

$$\sum_{i=1}^{N} R_i^N \le \kappa\nu \log(NP(N)) \triangleq C_a^N \text{ nats/channel use} \tag{38}$$

where R_i^N is the rate achievable from sensor i to the collector node, ν is a positive constant independent of N

$$\nu = \min\left(\frac{\epsilon}{1+2\epsilon}, \frac{1}{4}\right) \tag{39}$$

when N is large enough. Otherwise, the sum rate is bounded by a nonnegative constant as $N \to \infty$.

A proof of Theorem 5.1 is provided in Appendix I-C. Theorem 5.1 shows that when the sum power constraint is such that (37) is satisfied, the achievable rate increases with N. Furthermore, the achievable rate is the same as the upper bound on the achievable sum rate in (26) order-wise. Otherwise, the achievable sum rate is either a positive constant or decreases to zero, which will result in poor estimation performance at the collector node. The achievability scheme proposed in the proof of Theorem 5.1 incurs a delay that is proportional to the number of sensor nodes. From a practical point of view, it is desirable to have achievability schemes that perform better in terms of the latency. In this paper, we propose *an* achievability scheme that meets the lower bound order-wise, and leave the issue of developing better achievability schemes to future work.

Since the achievable rate region developed above is only characterized by the sum rate constraint, in the source coding part, for a fixed distortion constraint, we only need to characterize the achievable sum rate, rather than the achievable rate region. We apply [16, Theorem 1], generalized to N sensor nodes in [27, Theorem 1], to obtain an achievable sum rate–distortion point. The achievability scheme is an indirect version of the achievability scheme developed by Berger and Tung [28].

Theorem 5.2: For all Gaussian random processes, the following sum rate and distortion are achievable:

$$D_a^N(\theta) = \frac{1}{U_0}\int_0^{U_0}\left(K(u,u) - \frac{U_0}{N}\boldsymbol{\rho}_N^T(u) \times \left(\boldsymbol{\Sigma}_{\boldsymbol{N}}' + \theta\boldsymbol{I}\right)^{-1}\boldsymbol{\rho}_N(u)\right)du \tag{40}$$

$$R_a^N(\theta) = \sum_{k=0}^{N-1}\frac{1}{2}\log\left(1 + \frac{\mu_k^{(N)'}}{\theta}\right) \tag{41}$$

where $\boldsymbol{\rho}_N(u)$ is a column vector of size N, with its lth element being $K\left(u, \frac{(l-1)U_0}{N}\right)$, $\boldsymbol{\Sigma}_{\boldsymbol{N}}$ is defined as

$$\boldsymbol{\Sigma}_{\boldsymbol{N}} = E\left[\boldsymbol{S}_N\boldsymbol{S}_N^T\right] \tag{42}$$

i.e., the (i,j)th element of $\mathbf{\Sigma_N}$ is $K\left(\frac{(i-1)U_0}{N}, \frac{(j-1)U_0}{N}\right)$, $\mathbf{\Sigma'_N} = \frac{U_0}{N}\mathbf{\Sigma_N}$ and $\mu_0^{(N)'}, \mu_1^{(N)'}, \ldots, \mu_{N-1}^{(N)'}$ are the eigenvalues of $\mathbf{\Sigma'_N}$.

A proof of Theorem 5.2 is provided in Appendix I-D.

We further evaluate $D_a^N(\theta)$ in the next lemma.

Lemma 5.1: For all Gaussian random processes, we have

$$D_a^N(\theta) \le 2A^{(N)} + B^{(N)} + D_b^N(\theta) \tag{43}$$

where $A^{(N)}$, $B^{(N)}$, and $D_b^N(\theta)$ are defined as

$$A^{(N)} = \frac{1}{U_0}\sum_{i=1}^{N}\int_{\frac{i-1}{N}U_0}^{\frac{i}{N}U_0}\left(K(u,u) - K\left(\frac{i-1}{N}U_0, \frac{i-1}{N}U_0\right)\right)du + \frac{2}{U_0}\sum_{i=1}^{N}\int_{\frac{i-1}{N}U_0}^{\frac{i}{N}U_0}\left(\boldsymbol{\rho}_N\left(\frac{i-1}{N}U_0\right) - \boldsymbol{\rho}_N(u)\right)_i du \tag{44}$$

and

$$B^{(N)} = \frac{2}{U_0}\sum_{i=1}^{N}\int_{\frac{i-1}{N}U_0}^{\frac{i}{N}U_0}\left\|\boldsymbol{\rho}_N\left(\frac{i-1}{N}U_0\right) - \boldsymbol{\rho}_N(u)\right\| du \tag{45}$$

and

$$D_b^N(\theta) = \frac{1}{U_0}\sum_{k=0}^{N-1}\left(\frac{1}{\theta} + \frac{1}{\mu_k^{(N)'}}\right)^{-1} \tag{46}$$

respectively.

A proof of Lemma 5.1 is provided in Appendix I-E. Lemma 5.1 tells us that the expected distortion achieved by using the separation-based scheme is upper bounded by the sum of three types of distortion. The first two types of distortion, $A^{(N)}$ and $B^{(N)}$, have nothing to do with the rate and only depend on how well the samples approximate the entire random process. The third distortion, $D_b^N(\theta)$, depends on the rate through variable θ.

The function $R_a^N(\theta)$ is a strictly decreasing function of θ, thus, the inverse function exists, which we will denote as $\theta_a^N(R)$. Let us define $D_a(R)$ as the composition of the two functions $D_a^N(\theta)$ and $\theta_a^N(R)$, i.e.,

$$D_a(R) = D_a^N\left(\theta_a^N(R)\right). \tag{47}$$

An upper bound on the minimum achievable distortion, i.e., the achievable distortion by the separation-based scheme described above, is

$$D_u^N = D_a\left(C_a^N\right) \tag{48}$$

where C_a^N is defined in (38).

We will perform this calculation when the underlying random process is in $\mathcal{A}$.

B. The Class of Gaussian Random Processes in $\mathcal{A}$

We analyze the three types of distortion in (43) for Gaussian random processes in $\mathcal{A}$. We will focus on $A^{(N)}$ and $B^{(N)}$ in Lemma 5.2, and on $D_b^N(\theta)$ in Lemma 5.3.

Lemma 5.2: For Gaussian random processes in $\mathcal{A}$, we have

$$A^{(N)} = O\left(N^{-\alpha}\right) \tag{49}$$

$$B^{(N)} = O\left(N^{\frac{1}{2}-\alpha}\right). \tag{50}$$

A proof of Lemma 5.2 is provided in Appendix I-F. The result depends crucially on Condition 2 in the definition of $\mathcal{A}$ in Section III, i.e., the smoothness of the autocorrelation function $K(u,v)$. Note that since $1/2 < \alpha \le 1$, both $A^{(N)}$ and $B^{(N)}$ decrease to zero inverse polynomially as N goes to infinity.

It remains to calculate the functions $R_a^N(\theta)$ and $D_b^N(\theta)$ for random processes in $\mathcal{A}$. To do so, we need some properties of $\{\mu_k^{(N)'}\}_{k=0}^{N-1}$ which are stated in Lemmas A.3 and A.4 and proved in Appendix I-G. Lemma A.3 is of great importance, as it serves as a tool to link $\{\mu_k^{(N)'}\}_{k=0}^{N-1}$ to $\{\lambda_k\}_{k=0}^{\infty}$, which is used in the derivation of the lower bound in Section IV, through the lower and upper bounds $\{\lambda'_k\}_{k=0}^{\infty}$ and $\{\lambda''_k\}_{k=0}^{\infty}$. Armed with the properties of $\mu_k^{(N)'}$, λ'_k, and λ''_k in Lemmas A.1–A.4 in Appendices I-A and I-G, we can show the following lemma. First, we define two sequences ϑ_L^N and ϑ_U^N, which are functions of N, that satisfy

$$\lim_{N\to\infty}\frac{1}{\vartheta_L^N N^{\min\left(\frac{x\gamma}{2\tau}, \frac{\alpha x}{x-1}, \frac{\beta x}{x+\tau+1}\right)}} = 0, \quad \lim_{N\to\infty}\vartheta_U^N = 0 \tag{51}$$

Lemma 5.3: For Gaussian random processes in $\mathcal{A}$, for any constant $0 < \kappa < 1$, lower and upper bounds for the function $R_a^N(\theta)$ are

$$\frac{\kappa x d_l^{\frac{1}{x}}}{4}\theta^{-\frac{1}{x}} \le R_a^N(\theta) \le \frac{d_u^{\frac{1}{x}}\left(x^2 - (1-\log 2)x + (1-\log 2)\right)}{2(x-1)\kappa^2}\theta^{-\frac{1}{x}} \tag{52}$$

and an upper bound for the function $D_b^N(\theta)$ is

$$D_b^N(\theta) \le \frac{d_u^{\frac{1}{x}}\left(1 + \kappa^2(x-1)\right)}{\kappa^3(x-1)U_0}\theta^{1-\frac{1}{x}} \tag{53}$$

for $\theta \in \left[\vartheta_L^N, \vartheta_U^N\right]$ and N large enough.

A proof of Lemma 5.3 is provided in Appendix I-H. The proof of Lemma 5.3 uses Conditions 1, 2, and 3 in Section III. Let us define a sequence ϑ_{LL}^N, which is a function of N, that satisfies

$$\lim_{N\to\infty}\frac{1}{\vartheta_{LL}^N N^{\min\left(\frac{x\gamma}{2\tau}, \frac{(\alpha-1/2)x}{x-1}, \frac{\beta x}{x+\tau+1}\right)}} = 0. \tag{54}$$

Combining (43), (49), (50), (52), and (53), we have the following theorem.

Theorem 5.3: For Gaussian random processes in $\mathcal{A}$, for any constant $0 < \kappa < 1$, the achievable distortion–rate function $D_a(R)$ is upper bounded shown in (55) at the top of the following page, for R in the interval of

$$\left[\frac{d_u^{\frac{1}{x}}\left(x^2 - (1-\log 2)x + (1-\log 2)\right)}{2(x-1)\kappa^2}\left(\vartheta_U^N\right)^{-\frac{1}{x}}, \frac{\kappa x d_l^{\frac{1}{x}}}{4}\left(\vartheta_{LL}^N\right)^{-\frac{1}{x}}\right] \tag{56}$$

when N is large enough.

$$D_a(R) \leq \frac{d_u(1+\kappa^2(x-1))\left(x^2-(1-\log 2)x+(1-\log 2)\right)^{x-1}}{U_0\kappa^{2x+2}2^{x-1}(x-1)^x}R^{1-x} \tag{55}$$

A proof of Theorem 5.3 is provided in Appendix I-I. This theorem shows that when R is in the interval (56), the achievable distortion–rate function is the same as the lower bound on the distortion–rate function in (31) order-wise.

Theorem 5.4: For Gaussian random processes in $\mathcal{A}$, when the sum power constraint satisfies (37) and

$$\lim_{N\to\infty} \frac{NP(N)}{e^{N^{\min\left(\frac{\gamma}{2\tau},\frac{2\alpha-1}{2(x-1)},\frac{\beta}{x+\tau+1}\right)}}} = 0 \tag{57}$$

an upper bound on the minimum achievable expected distortion, or equivalently, the achievable rate in the separation-based scheme, is as shown in (58) and (59) at the bottom of the page, when N is large enough.

A proof of Theorem 5.4 is provided in Appendix I-J. Theorem 5.4 implies that, when the sum power constraint satisfies (37) and (57), the minimum achievable expected distortion is

$$O\left(\left(\frac{1}{\log(NP(N))}\right)^{x-1}\right). \tag{60}$$

For the interesting cases of $P(N) = NP_{\text{ind}}$ and $P(N) = P_{\text{tot}}$, the upper bound on the minimum achievable expected distortion decays to zero at the rate of

$$\frac{1}{(\log N)^{x-1}}. \tag{61}$$

When the sum power constraint is such that (37) is not satisfied, an upper bound on the minimum achievable expected distortion is $\Theta(1)$.

VI. Comparison of the Lower and Upper Bounds for Gaussian Random Processes in $\mathcal{A}$

A. Order-Wise Comparison of Lower and Upper Bounds

In this section, we compare the lower bound in (35) and the upper bound in (60). When the sum power constraint is large, i.e., $P(N)$ is so large that (57) is not satisfied, our methods in finding the upper bound do not apply. Even though our lower bound in (35) is valid, we have not shown whether the lower and upper bounds meet. However, in this case, $P(N)$ is larger than $\frac{e^{N^{\min\left(\frac{\gamma}{2\tau},\frac{2\alpha-1}{2(x-1)},\frac{\beta}{x+\tau+1}\right)}}}{N}$, and this region of sum power constraint is not of practical interest.

When the sum power constraint is medium, i.e., $P(N)$ is in the wide range of $N^{-1/2+\epsilon}$ to $\frac{e^{N^{\min\left(\frac{\gamma}{2\tau},\frac{2\alpha-1}{2(x-1)},\frac{\beta}{x+\tau+1}\right)}}}{N}$, our lower and upper bounds do meet and the minimum achievable expected distortion is

$$D^N = \Theta\left(\left(\frac{1}{\log(NP(N))}\right)^{x-1}\right). \tag{62}$$

One possible order-optimal achievability scheme is a separation-based scheme, which uses distributed rate–distortion coding as described in [16] and optimal single-user channel coding with amplify-and-forward method as described in [17]. In fact, when the sum power constraint is medium, as shown in (31) and (55), lower and upper bounds on the distortion–rate function $D_p(R)$ and $D_a(R)$ coincide order-wise. In addition, as shown in (27) and (38), the lower and upper bounds on the achievable sum rate, C_a^N and C_u^N coincide order-wise as well. The practically interesting cases of $P(N) = NP_{\text{ind}}$ and $P(N) = P_{\text{tot}}$ fall into this region of medium sum power constraint. In both cases, the minimum achievable expected distortion decreases to zero at the rate of

$$\frac{1}{(\log N)^{x-1}}. \tag{63}$$

Hence, the sum power constraint $P(N) = P_{\text{tot}}$ performs as well as $P(N) = NP_{\text{ind}}$ "order-wise," and therefore, in practice we may prefer to choose $P(N) = P_{\text{tot}}$. In fact, we can decrease the sum power constraint to $P(N) = \Theta\left(N^{-1/3}\right)$ and the minimum achievable distortion will still decrease to zero at the rate in (63).

When the sum power constraint is small, i.e., $P(N)$ ranges from N^{-1} to $N^{-1/2}$, our lower and upper bounds do not meet. Our lower bound in (35) decreases to zero as $\frac{1}{(\log N)^{x-1}}$ but our upper bound is a nonzero constant. The main discrepancy between our lower and upper bounds comes from the gap between the lower and upper bounds on the sum capacities C_a^N and C_u^N for a cooperative multiple-access channel. In fact, when the sum power constraint is small, as shown in (31) and (55), lower and upper bounds on the distortion–rate function $D_p(R)$ and $D_a(R)$ still coincide order-wise. This sum power constraint region should be of practical interest, because in this region, the sum power constraint is quite low, and yet the lower bound on the distortion is of the same order as one would obtain with any $P(N)$ which increases polynomially with N. Hence, from the lower bound, it seems that this region potentially has good performance. However, our separation-based upper bound does

$$D_u^N = D_a\left(C_a^N\right) \tag{58}$$

$$\leq \frac{d_u(1+\kappa^2(x-1))\left(x^2-(1-\log 2)x+(1-\log 2)\right)^{x-1}}{U_0\kappa^{3x+1}2^{x-1}(x-1)^x\nu^{x-1}}\left(\frac{1}{\log(NP(N))}\right)^{x-1} \tag{59}$$

not meet the lower bound, and whether the lower bound can be achieved remains an open problem.

When the sum power constraint is very small, i.e., $P(N)$ is less than N^{-1}, our lower and upper bounds meet and the minimum achievable expected distortion is a constant that does not decrease to zero with increasing N. This case is not of practical interest because of the unacceptable distortion.

In the case of Gauss–Markov random process, we have $x = 2$ and $\alpha = \beta = \tau = \gamma = 1$. Inserting these values into the above results, we see that in the medium sum power constraint region, i.e., $P(N)$ is in the wide range of $N^{-1/2+\epsilon}$ to $\frac{e^{N^{1/4}}}{N}$, the minimum achievable expected distortion is

$$D^N = \Theta\left(\frac{1}{\log(NP(N))}\right). \tag{64}$$

For the Gauss–Markov random process, in the cases of $P(N) = NP_{\text{ind}}$ and $P(N) = P_{\text{tot}}$, the minimum achievable expected distortion decreases to zero at the rate of

$$\frac{1}{\log N}. \tag{65}$$

The conclusions in (64) and (65) were derived in [1] under a different channel assumption. For the channel assumption in [1], the order-optimal achievability scheme was determined to be a decode-and-forward based scheme. The range of medium power constraints was shown to be slightly larger in [1], i.e., $P(N)$ in the range of $N^{-1/2+\epsilon}$ to $\frac{e^{N^{1/3}}}{N}$, and this is because it was specifically derived for the Gauss–Markov process, instead of general Gaussian random processes as in this work.

B. Comparison of the Constants in the Lower and Upper Bounds

Though the lower and upper bounds meet order-wise in a wide range of sum power constraints, the constants in front of them are different and we aim to compare these constants for various sum power constraints in this section.

Combining (33) and (59), when $P(N)$ satisfies (37) and (57), the minimum distortion D^N satisfies (66), shown at the bottom of the page. Note that κ can be made as close to 1 as possible for large enough N. Let $\pi(x,\nu)$ be the ratio of the constant in the lower bound and the constant in the upper bound when N is large enough. Then

$$\pi(x,\nu) = \frac{d_l}{d_u}(2\nu)^{x-1}\left(\frac{x^2-x}{x^2-(1-\log 2)x+(1-\log 2)}\right)^{x-1}. \tag{67}$$

Here, x is a parameter of the underlying Gaussian random process and ν depends on the sum power constraint of the sensor nodes $P(N)$. It is straightforward to see that since from (39), $\nu \le 1/4$, $\pi(x,\nu)$ is a monotonically decreasing function of x for a fixed ν. Hence, we conclude that the constants in front of the lower and upper bounds differ more as x gets large. Since x is an indication of how concentrated the randomness of the random process is, this means that the more evenly distributed the randomness, the more the constants in the lower and upper bounds meet. For a fixed underlying random process, i.e., for a fixed x, $\pi(x,\nu)$ is a decreasing function of ν. This means that the less the sum power constraint we have, the more different the constants will be.

In the Gauss–Markov random process $x = 2$ and $d_l = d_u$. When $P(N) = NP_{\text{ind}}$ and $P(N) = P_{\text{tot}}$, the ratio of the two constants is

$$\pi(2,1/4) = \frac{1}{3+\log 2} \simeq 0.2708. \tag{68}$$

When $P(N) = \Theta(N^{-\omega})$, $0 < \omega < \frac{1}{2}$, the ratio of the two constants is

$$\pi\left(2,\frac{1}{2}-\frac{1}{4}\frac{1}{1-\omega}\right) = \left(\frac{1}{2}-\frac{1}{4}\frac{1}{1-\omega}\right)\frac{4}{3+\log 2}. \tag{69}$$

For example, when $P(N) = \Theta\left(N^{-1/3}\right)$, the ratio of the constants is

$$\pi(2,1/8) = \frac{1}{2}\pi(2,1/4) \simeq 0.1354 \tag{70}$$

VII. Further Remarks

We have shown that the minimum achievable expected distortion behaves order-wise as

$$\Theta\left(\left(\frac{1}{\log(NP(N))}\right)^{x-1}\right). \tag{71}$$

Due to the order-optimality of separation, this result can be generalized straightforwardly to several other scenarios.

The result in (71) still holds when we allow the collector node to use its transmit antenna with an arbitrary power constraint. The collector node, using its transmit antenna, can send some form of feedback to the sensor nodes. However, the lower bound on the minimum distortion remains unchanged in this case, because in deriving our lower bound, we assumed that all sensor nodes know the entire random process, thus, forming a

$$\kappa^2\left(1+\frac{\kappa}{x-1}\right)(\kappa x)^{x-1}\frac{d_l}{U_0}\left(\frac{1}{\log(NP(N))}\right)^{x-1} \le D^N \le \frac{d_u(1+\kappa^2(x-1))\left(x^2-(1-\log 2)x+(1-\log 2)\right)^{x-1}}{U_0\kappa^{3x+1}2^{x-1}(x-1)^x\nu^{x-1}}\left(\frac{1}{\log(NP(N))}\right)^{x-1}. \tag{66}$$

point-to-point system. In a point-to-point system, feedback, perfect or not, does not change the capacity. Meanwhile, our upper bound is still valid, as in this achievable scheme, we choose not to utilize the feedback link. Hence, our result in (71) remains valid.

The result in (71) still holds when we allow K channel uses per realization of the random process, where K is a constant independent of N. This is because both lower and upper bounds are derived using separation-based schemes. The minimum achievable distortion still behaves as (71), and the number K will only affect the constant in front. Due to the same reasoning, the minimum achievable distortion behaves as (71) when we allow multiple transmit and receive antennas at each node, as long as the number of antennas on each node is a constant, independent of N.

The assumption of the polynomial decay of the eigenvalues plays a key role for the separation principle to hold order-wise. For example, when the eigenvalues decrease exponentially, i.e., the kth eigenvalue is roughly e^{-k}, the distortion rate function of the lower bound is

$$D_p(R) = \Theta\left(\sqrt{R}e^{-2\sqrt{R}}\right). \tag{72}$$

Thus, in the lower bound, the distortion goes to zero almost exponentially with the rate R, as opposed to the polynomial decrease in R as in (31). It can be shown, using the exact same proof techniques as those used in Section V-B, that the achievable distortion–rate function is the same order as (72), for a wide range of sum power constraints. However, in the channel coding part, the converse and the achievability of the sum rate meet only order-wise, i.e., the lower and upper bounds on the sum rate are of the form $c_1 \log(NP(N))$ and $c_2 \log(NP(N))$ where $c_1 < c_2$. The difference in the constants in the lower and upper bounds on the sum rate will cause an order difference in the distortion, i.e., $\sqrt{c_1 \log(NP(N))}e^{-2\sqrt{c_1 \log(NP(N))}}$ is strictly of a larger order than $\sqrt{c_2 \log(NP(N))}e^{-2\sqrt{c_2 \log(NP(N))}}$ for $c_1 < c_2$. Hence, when the underlying random process is such that the eigenvalues decrease exponentially, the separation principle does not hold, even order-wise. This agrees with the observation made in Section VI-B that the constants in front of the lower and upper bounds differ more as x gets large.

For simplicity, we have considered only one-dimensional spatial random processes. We expect the generalization to two-dimensional random fields to be straightforward, but nonetheless tedious. Our results do not generalize straightforwardly when the samples that the sensor nodes obtain are subject to noise. Since the lower bound of assuming all sensors know the entire random process would remain the same with or without noise, the lower bound becomes too loose. Hence, the optimal performance under the noisy sensor scenario remains open.

VIII. Conclusion

In this paper, we investigated the performance of dense sensor networks by studying the joint source–channel coding problem. We provided separation-based lower and upper bounds for the minimum achievable expected distortion when the underlying random process is Gaussian. When the random process satisfies some general conditions, such as polynomial decrease rate of the ordered eigenvalues of the random process, i.e., the kth eigenvalue is roughly k^{-x}, we evaluated the lower and upper bounds explicitly, and showed that they are both of order $\frac{1}{(\log(NP(N)))^{x-1}}$ for a wide range of sum power constraints ranging from $N^{-\frac{1}{2}+\epsilon}$ to $\frac{e^{N^{\min\left(\frac{\gamma}{2\tau}, \frac{2\alpha-1}{2(x-1)}, \frac{\beta}{x+\tau+1}\right)}}}{N}$. In the most interesting cases, when the sum power constraint is a constant or grows linearly with N, the minimum achievable expected distortion decreases to zero at the rate of $\frac{1}{(\log N)^{x-1}}$. For random processes that satisfy these general conditions, under these power constraints, we have found an order-optimal scheme that is separation-based, and is composed of distributed rate–distortion coding [16] and amplify-and-forward channel coding [17].

Appendix I

A. Some Properties of λ'_k and λ''_k

In this subsection, we provide two lemmas which characterize some properties of $\{\lambda'_k\}_{k=0}^{\infty}$ and $\{\lambda''_k\}_{k=0}^{\infty}$, defined in (12) and (13), which will be useful in deriving our main results.

Lemma A.1: For any constant $0 < \kappa < 1$, we have

$$\sum_{k=\left\lfloor \left(\frac{d_l}{\theta}\right)^{\frac{1}{x}} - c_l + 1 \right\rfloor}^{\infty} \lambda'_k \geq \frac{\kappa d_l^{\frac{1}{x}}}{(x-1)} \theta^{1-\frac{1}{x}} \tag{73}$$

and

$$\sum_{k=0}^{\left\lfloor \left(\frac{d_l}{\theta}\right)^{\frac{1}{x}} - c_l \right\rfloor} \frac{1}{2} \log\left(\frac{\lambda'_k}{\theta}\right) \geq \frac{\kappa x d_l^{\frac{1}{x}}}{2} \theta^{-\frac{1}{x}} \tag{74}$$

when θ is small enough.

Lemma A.2: For any constant $0 < \kappa < 1$, we have

$$\sum_{k=\left\lfloor \left(\frac{d_u}{\theta}\right)^{\frac{1}{x}} + c_u \right\rfloor + 1}^{\infty} \lambda''_k \leq \frac{d_u^{\frac{1}{x}}}{(x-1)\kappa} \theta^{1-\frac{1}{x}} \tag{75}$$

and

$$\sum_{k=0}^{\left\lfloor \left(\frac{d_u}{\theta}\right)^{\frac{1}{x}} + c_u \right\rfloor} \frac{1}{2} \log\left(1 + \frac{\lambda''_k}{\theta}\right) \leq \left(\frac{\log 2 + x}{2\kappa}\right) d_u^{\frac{1}{x}} \theta^{-\frac{1}{x}} \tag{76}$$

when θ is small enough.

1) Proof of Lemma A.1: We will first prove (73).

$$\sum_{k=\left\lfloor \left(\frac{d_l}{\theta}\right)^{\frac{1}{x}} - c_l + 1 \right\rfloor}^{\infty} \lambda'_k = \sum_{k=\left\lfloor \left(\frac{d_l}{\theta}\right)^{\frac{1}{x}} - c_l + 1 \right\rfloor}^{\infty} \frac{d_l}{(k+c_l)^x} \tag{77}$$

$$\geq \frac{d_l}{x-1} \frac{1}{\left\lfloor \left(\frac{d_l}{\theta}\right)^{\frac{1}{x}} + 1 \right\rfloor^{x-1}} \tag{78}$$

$$\geq \frac{\kappa d_l^{\frac{1}{x}}}{(x-1)} \theta^{1-\frac{1}{x}} \tag{79}$$

where (77) is true when θ is small enough, more specifically, when $\left\lfloor \left(\frac{d_l}{\theta}\right)^{\frac{1}{x}} - c_l + 1 \right\rfloor > K_0$. We have (78) because of the inequality

$$\sum_{k=n}^{\infty} \frac{1}{k^x} \geq \int_n^{\infty} \frac{1}{y^x} dy = \frac{1}{(x-1)n^{x-1}} \tag{80}$$

and (79) is true when θ is small enough, i.e., for any $0 < \kappa < 1$, there exists a $\theta_0(\kappa) > 0$ such that when $0 < \theta \leq \theta_0(\kappa)$, (79) is true.

Next, we will prove (74). See (81)–(85) at the bottom of the page, where (81) is true when θ is small enough, more specifically, when $\left\lfloor \left(\frac{d_l}{\theta}\right)^{\frac{1}{x}} - c_l \right\rfloor > K_0$, and (83) follows by using Stirling's approximation

$$n! < \sqrt{2\pi} n^{n+\frac{1}{2}} e^{-n+\frac{1}{12n}}. \tag{86}$$

Equation (84) follows because c_3 is a constant, independent of θ, defined as

$$c_3 \triangleq \sum_{k=0}^{K_0} \frac{1}{2} \log\left(\frac{\lambda_k}{d_l}\right) + \frac{x}{2} \log((K_0 + c_l)!) - \frac{x}{4} \log(2\pi) \tag{87}$$

and (85) is true when θ is small enough, i.e., for any $0 < \kappa < 1$, there exists a $\theta_1(\kappa) > 0$ such that when $0 < \theta \leq \theta_1(\kappa)$, (85) is true.

Therefore, for any $0 < \kappa < 1$, (73) and (74) hold when θ is small enough.

2) Proof of Lemma A.2: We will first prove (75).

$$\sum_{k=\left\lfloor \left(\frac{d_u}{\theta}\right)^{\frac{1}{x}} + c_u \right\rfloor + 1}^{\infty} \lambda_k'' = \sum_{k=\left\lfloor \left(\frac{d_u}{\theta}\right)^{\frac{1}{x}} + c_u \right\rfloor + 1}^{\infty} \frac{d_u}{(k - c_u)^x} \tag{88}$$

$$= \frac{d_u}{(x-1)\left(\left\lfloor \left(\frac{d_u}{\theta}\right)^{\frac{1}{x}} \right\rfloor\right)^{x-1}} \tag{89}$$

$$\leq \frac{d_u^{\frac{1}{x}}}{(x-1)\kappa} \theta^{1-\frac{1}{x}} \tag{90}$$

where (88) follows when θ is small enough, more specifically, when $\left\lfloor \left(\frac{d_u}{\theta}\right)^{\frac{1}{x}} + c_u \right\rfloor + 1 > K_0$. In obtaining (89) we used

$$\sum_{k=n}^{\infty} \frac{1}{k^x} \leq \int_{n-1}^{\infty} \frac{1}{y^x} dy = \frac{1}{(x-1)(n-1)^{x-1}} \tag{91}$$

and (90) follows when θ is small enough, i.e., for any $0 < \kappa < 1$, there exists a $\theta_2(\kappa) > 0$ such that when $0 < \theta \leq \theta_2(\kappa)$, (90) is true.

Next, we will prove (76). See (92)–(97) at the top of the following page, where (92) is true when θ is small enough, more specifically, when $\left\lfloor \left(\frac{d_u}{\theta}\right)^{\frac{1}{x}} + c_u \right\rfloor > K_0$. We have (93) because

$$\frac{d_u}{(k - c_u)^x \theta} > 1 \tag{98}$$

for all k between $K_0 + 1$ and $\left\lfloor \left(\frac{d_u}{\theta}\right)^{\frac{1}{x}} + c_u \right\rfloor$, and when θ is small enough such that

$$\theta \leq \lambda_k, \qquad k = 1, 2, \ldots, K_0. \tag{99}$$

We have (95) because we defined

$$c_1 \triangleq \sum_{k=1}^{K_0} \frac{1}{2} \log \frac{2\lambda_k}{d_u} - \sum_{k=c_u+1}^{K_0} \frac{1}{2} \log \frac{2}{(k - c_u)^x}. \tag{100}$$

$$\sum_{k=0}^{\left\lfloor \left(\frac{d_l}{\theta}\right)^{\frac{1}{x}} - c_l \right\rfloor} \frac{1}{2} \log\left(\frac{\lambda_k'}{\theta}\right) = \sum_{k=0}^{K_0} \frac{1}{2} \log\left(\frac{\lambda_k}{\theta}\right) + \sum_{k=K_0+1}^{\left\lfloor \left(\frac{d_l}{\theta}\right)^{\frac{1}{x}} \right\rfloor - c_l} \frac{1}{2} \log\left(\frac{d_l}{(k + c_l)^x \theta}\right) \tag{81}$$

$$= \sum_{k=0}^{K_0} \frac{1}{2} \log\left(\frac{\lambda_k}{d_l}\right) + \frac{1}{2}\left(\left\lfloor \left(\frac{d_l}{\theta}\right)^{\frac{1}{x}} \right\rfloor - c_l + 1\right) \log\left(\frac{d_l}{\theta}\right) - \frac{x}{2} \log\left(\left\lfloor \left(\frac{d_l}{\theta}\right)^{\frac{1}{x}} \right\rfloor !\right) + \frac{x}{2} \log((K_0 + c_l)!) \tag{82}$$

$$\geq \frac{1}{2}\left(\left\lfloor \left(\frac{d_l}{\theta}\right)^{\frac{1}{x}} \right\rfloor - c_l + 1\right) \log\left(\frac{d_l}{\theta}\right) - \frac{x}{2}\left(\left\lfloor \left(\frac{d_l}{\theta}\right)^{\frac{1}{x}} \right\rfloor + \frac{1}{2}\right) \log\left\lfloor \left(\frac{d_l}{\theta}\right)^{\frac{1}{x}} \right\rfloor + \frac{x}{2}\left\lfloor \left(\frac{d_l}{\theta}\right)^{\frac{1}{x}} \right\rfloor - \frac{x}{24\left\lfloor \left(\frac{d_l}{\theta}\right)^{\frac{1}{x}} \right\rfloor} + \sum_{k=0}^{K_0} \frac{1}{2} \log\left(\frac{\lambda_k}{d_l}\right) + \frac{x}{2} \log((K_0 + c_l)!) - \frac{x}{4} \log(2\pi) \tag{83}$$

$$\geq \frac{x}{2}\left\lfloor \left(\frac{d_l}{\theta}\right)^{\frac{1}{x}} \right\rfloor + \frac{x}{2}\left(-c_l + \frac{1}{2}\right) \log\left\lfloor \left(\frac{d_l}{\theta}\right)^{\frac{1}{x}} \right\rfloor - \frac{x}{24\left\lfloor \left(\frac{d_l}{\theta}\right)^{\frac{1}{x}} \right\rfloor} + c_3 \tag{84}$$

$$\geq \frac{\kappa x d_l^{\frac{1}{x}}}{2} \theta^{-\frac{1}{x}} \tag{85}$$

$$\sum_{k=0}^{\left\lfloor \left(\frac{d_u}{\theta}\right)^{\frac{1}{x}}+c_u \right\rfloor} \frac{1}{2}\log\left(1+\frac{\lambda_k''}{\theta}\right) = \sum_{k=0}^{K_0}\frac{1}{2}\log\left(1+\frac{\lambda_k}{\theta}\right) + \sum_{k=K_0+1}^{\left\lfloor \left(\frac{d_u}{\theta}\right)^{\frac{1}{x}}+c_u \right\rfloor} \frac{1}{2}\log\left(1+\frac{d_u}{(k-c_u)^x\theta}\right) \tag{92}$$

$$\leq \sum_{k=0}^{K_0}\frac{1}{2}\log\left(\frac{2\lambda_k}{\theta}\right) + \sum_{k=K_0+1}^{\left\lfloor \left(\frac{d_u}{\theta}\right)^{\frac{1}{x}}+c_u \right\rfloor} \frac{1}{2}\log\left(\frac{2d_u}{(k-c_u)^x\theta}\right) \tag{93}$$

$$= \sum_{k=0}^{K_0}\frac{1}{2}\log\left(\frac{2\lambda_k}{\theta}\right) + \sum_{k=c_u+1}^{\left\lfloor \left(\frac{d_u}{\theta}\right)^{\frac{1}{x}}+c_u \right\rfloor} \frac{1}{2}\log\left(\frac{2d_u}{(k-c_u)^x\theta}\right) - \sum_{k=c_u+1}^{K_0}\frac{1}{2}\log\left(\frac{2d_u}{(k-c_u)^x\theta}\right) \tag{94}$$

$$= \left\lfloor \left(\frac{d_u}{\theta}\right)^{\frac{1}{x}} \right\rfloor \frac{1}{2}\log 2 - \frac{x}{2}\log\left(\left\lfloor \left(\frac{d_u}{\theta}\right)^{\frac{1}{x}} \right\rfloor !\right) + \frac{1}{2}\left(\left\lfloor \left(\frac{d_u}{\theta}\right)^{\frac{1}{x}} \right\rfloor\right)\log\frac{d_u}{\theta} + \frac{c_u+1}{2}\log\frac{d_u}{\theta} + c_1 \tag{95}$$

$$\leq \left\lfloor \left(\frac{d_u}{\theta}\right)^{\frac{1}{x}} \right\rfloor \frac{1}{2}\log 2 - \frac{x}{4}\log(2\pi) - \frac{x}{2}\left(\left\lfloor \left(\frac{d_u}{\theta}\right)^{\frac{1}{x}} \right\rfloor + \frac{1}{2}\right)\log\left\lfloor \left(\frac{d_u}{\theta}\right)^{\frac{1}{x}} \right\rfloor + \frac{x}{2}\left\lfloor \left(\frac{d_u}{\theta}\right)^{\frac{1}{x}} \right\rfloor - \frac{x}{24\left\lfloor \left(\frac{d_u}{\theta}\right)^{\frac{1}{x}} \right\rfloor + 2} + \frac{1}{2}\left(\left\lfloor \left(\frac{d_u}{\theta}\right)^{\frac{1}{x}} \right\rfloor\right)\log\frac{d_u}{\theta} + \frac{c_u+1}{2}\log\frac{d_u}{\theta} + c_1 \tag{96}$$

$$\leq \left(\frac{\log 2 + x}{2\kappa}\right) d_u^{\frac{1}{x}}\theta^{-\frac{1}{x}} \tag{97}$$

We used Stirling's approximation

$$n! > \sqrt{2\pi}n^{n+\frac{1}{2}}e^{-n+\frac{1}{12n+1}} \tag{101}$$

to obtain (96), and (97) follows when θ is small enough, i.e., for any $0 < \kappa < 1$, there exists a $\theta_3(\kappa) > 0$ such that when $0 < \theta \leq \theta_3(\kappa)$, (97) is true.

Therefore, for any $0 < \kappa < 1$, (75) and (76) hold when θ is small enough.

B. Proof of Lemma 4.1

For any $0 < \kappa < 1$, when θ is small enough, the results of Lemma A.1 hold.

From (24), we have

$$R(\theta) = \sum_{k=0}^{\infty}\max\left(0, \frac{1}{2}\log\left(\frac{\lambda_k}{\theta}\right)\right) \tag{102}$$

$$\geq \sum_{k=0}^{\infty}\max\left(0, \frac{1}{2}\log\left(\frac{\lambda_k'}{\theta}\right)\right) \tag{103}$$

$$= \sum_{k=0}^{\left\lfloor \left(\frac{d_l}{\theta}\right)^{\frac{1}{x}}-c_l \right\rfloor} \frac{1}{2}\log\left(\frac{\lambda_k'}{\theta}\right) \tag{104}$$

$$\geq \frac{\kappa x d_l^{\frac{1}{x}}}{2}\theta^{-\frac{1}{x}} \tag{105}$$

where in (103) we have used the definition of sequence λ_k' in (12) and the observation in (14). Equation (104) follows when θ is small enough, more specifically, when $\theta < \lambda_{K_0}$ and $\left\lfloor \left(\frac{d_l}{\theta}\right)^{\frac{1}{x}} - c_l \right\rfloor > K_0$. Equation (105) follows from (74) in Lemma A.1.

From (25), we have

$$D(\theta) = U_0^{-1}\sum_{k=0}^{\infty}\min(\theta, \lambda_k) \tag{106}$$

$$\geq U_0^{-1}\sum_{k=0}^{\infty}\min(\theta, \lambda_k') \tag{107}$$

$$= U_0^{-1}\sum_{k=0}^{\left\lfloor \left(\frac{d_l}{\theta}\right)^{\frac{1}{x}}-c_l \right\rfloor}\theta + U_0^{-1}\sum_{\left\lfloor \left(\frac{d_l}{\theta}\right)^{\frac{1}{x}}-c_l+1 \right\rfloor}^{\infty}\lambda_k' \tag{108}$$

$$\geq U_0^{-1}\left(\left\lfloor \left(\frac{d_l}{\theta}\right)^{\frac{1}{x}} \right\rfloor - c_l + 1\right)\theta + U_0^{-1}\frac{\kappa d_l^{\frac{1}{x}}}{(x-1)}\theta^{1-\frac{1}{x}} \tag{109}$$

$$\geq \kappa\left(1+\frac{\kappa}{x-1}\right)\frac{d_l^{\frac{1}{x}}}{U_0}\theta^{1-\frac{1}{x}} \tag{110}$$

where in (107) we have used the definition of sequence λ_k' in (12) and the observation in (14). Equation (108) follows when θ is small enough, more specifically, when $\theta < \lambda_{K_0}$ and $\left\lfloor \left(\frac{d_l}{\theta}\right)^{\frac{1}{x}} - c_l + 1 \right\rfloor > K_0$. Equation (109) follows from (73) in Lemma A.1. Equation (110) is true for small enough θ, i.e., for any $0 < \kappa < 1$, there exists a $\theta_4(\kappa) > 0$ such that when $0 < \theta \leq \theta_4(\kappa)$, (110) is true.

Therefore, for any $0 < \kappa < 1$, (29) and (30) hold when θ is small enough.

C. Proof of Theorem 5.1

We will show that each sensor node i can achieve a rate of C_a^N, while the other sensor nodes have rate zero, then by the time sharing argument [29] we can achieve the rate region of

$$\sum_{i=1}^{N} R_i^N \leq C_a^N. \tag{111}$$

We will consider the transmission of the data of node i. All other sensor nodes have no data to transmit and are helping with the communication between sensor node i and the collector node. Node i codes its message using capacity-achieving single-user coding techniques with codeword length $\bar{n}$. Each codeword symbol requires two time slots. In the first time slot, node i transmits its codeword symbol using power $P(N)$. All other nodes remain silent, and receive a noisy version of node i's transmitted signal. The collector node ignores its received signal, which is suboptimal but eases calculation and does not affect the scaling law of the achievable rate. In the second time slot, all sensor nodes, except node i, amplify and forward what they have received in the previous time slot to the collector node using a sum power constraint $P(N)$. The collector node, after $2\bar{n}$ time slots, decodes using capacity-achieving single-user decoding techniques. The scheme described satisfies the sum power constraint of $P(N)$. Now, we calculate the rate achievable with this scheme. In the first time slot, sensor node j receives

$$Y_j = h_{ij}X_i + Z_j, \qquad i,j = 1,2,\ldots,N,\ j \neq i \tag{112}$$

and in the second time slot, sensor node j transmits

$$\begin{aligned} X_j &= \beta_{ij}Y_j && (113)\\ &= \beta_{ij}h_{ij}X_i + \beta_{ij}Z_j, \quad i,j = 1,2,\ldots,N,\ j \neq i && (114)\end{aligned}$$

where β_{ij} is the scaling coefficient of node j when it amplifies the signal it received from node i. In order to satisfy the sum power constraints, $\{\beta_{ij}\}_{j=1,j\neq i}^N$ have to satisfy

$$\sum_{j=1,j\neq i}^{N} \beta_{ij}^2 \left(h_{ij}^2 P(N) + 1\right) \leq P(N), \qquad i = 1,2,\ldots,N. \tag{115}$$

The collector node receives

$$\begin{aligned} Y_0 &= \sum_{j=1,j\neq i}^{N} h_{j0}X_j + Z_0 && (116)\\ &= \left(\sum_{j=1,j\neq i}^{N} \beta_{ij}h_{ij}h_{j0}\right) X_i + \left(\sum_{j=1,j\neq i}^{N} h_{j0}\beta_{ij}Z_j\right) + Z_0. && (117)\end{aligned}$$

Therefore, the achievable rate is

$$\frac{1}{4}\log\left(1 + \frac{\left(\sum_{j=1,j\neq i}^{N} \beta_{ij}h_{ij}h_{j0}\right)^2 P(N)}{\sum_{j=1,j\neq i}^{N} \left(\beta_{ij}h_{j0}\right)^2 + 1}\right) \tag{118}$$

where we have $1/4$ because we used two time slots to transmit one codeword symbol. We choose

$$\beta_{ij} = \zeta h_{ij}h_{j0} \tag{119}$$

where, in order to satisfy the power constraint, the constant ζ must satisfy

$$\zeta^2 \leq \frac{P(N)}{\left(\sum_{j=1,j\neq i}^{N} h_{ij}^4 h_{j0}^2\right) P(N) + \left(\sum_{j=1,j\neq i}^{N} h_{ij}^2 h_{j0}^2\right)}. \tag{120}$$

We can choose ζ as

$$\zeta^2 = \frac{P(N)}{\bar{h}_u^6 N P(N) + \bar{h}_u^4 N}. \tag{121}$$

Thus, from (118), a lower bound on the achievable rate is

$$\begin{aligned} &\frac{1}{4}\log\left(1 + \frac{\zeta^2\left(\sum_{j=1,j\neq i}^{N} h_{ij}^2 h_{j0}^2\right)^2 P(N)}{\zeta^2\left(\sum_{j=1,j\neq i}^{N} h_{ij}^2 h_{j0}^4\right) + 1}\right)\\ &\quad \geq \frac{1}{4}\log\left(1 + \frac{\bar{h}_l^8 \zeta^2 (N-1)^2 P(N)}{\bar{h}_u^6 \zeta^2 N + 1}\right) \triangleq C_b^N. \end{aligned} \tag{122}$$

Clearly, rate C_b^N can be achievable by any node i. We have

$$\begin{aligned} C_b^N &= \frac{1}{4}\log\left(1 + \frac{\bar{h}_l^8 \left(P(N)\right)^2 \frac{(N-1)^2}{N}}{2\bar{h}_u^6 P(N) + \bar{h}_u^4}\right) && (123)\\ &\geq \frac{1}{4}\log\left(1 + \frac{\bar{h}_l^8 \left(P(N)\right)^2 N}{4\bar{h}_u^6 P(N) + 2\bar{h}_u^4}\right) && (124)\end{aligned}$$

where the last step follows when N is large enough such that $\frac{(N-1)^2}{N} > \frac{N}{2}$.

When $P(N)$ is such that

$$\lim_{N\to\infty} \frac{1}{P(N)} = 0 \tag{125}$$

for any $0 < \kappa < 1$, we have

$$\begin{aligned} C_b^N &\geq \frac{1}{4}\log\left(1 + \frac{\bar{h}_l^8 \left(P(N)\right)^2 N}{8\bar{h}_u^6 P(N)}\right) && (126)\\ &= \frac{1}{4}\log\left(1 + \frac{\bar{h}_l^8}{8\bar{h}_u^6} N P(N)\right) && (127)\\ &\geq \frac{\kappa}{4}\log\left(N P(N)\right) && (128)\end{aligned}$$

for N large enough, i.e., there exists $N_1(\kappa) > 0$, such that when $N > N_1(\kappa)$, (126) and (128) are true.

When $P(N)$ is such that

$$\lim_{N\to\infty} P(N) = l \tag{129}$$

and l is a number that satisfies $0 < l < \infty$, fix some small $l_0 > 0$, there exists an $N_2(l_0) > 0$ such that when $N > N_2(l_0)$, we have

$$l - l_0 < P(N) < l + l_0. \tag{130}$$

Hence, when $N > N_2(l_0)$, for any $0 < \kappa < 1$

$$C_b^N \geq \frac{1}{4}\log\left(1 + \frac{\bar{h}_l^8(l-l_0)}{4\bar{h}_u^6(l+l_0)+2\bar{h}_u^4}P(N)N\right) \quad (131)$$

$$\geq \frac{\kappa}{4}\log\left(NP(N)\right) \quad (132)$$

where the last step follows when N is large enough, i.e., when there exists an $N_3(\kappa) > 0$, such that when $N > \max\left(N_2(l_0), N_3(\kappa)\right)$, (132) is true.

When $P(N)$ is such that

$$\lim_{N\to\infty} P(N) = 0 \quad (133)$$

and there exists a constant $0 < \epsilon < \frac{1}{2}$, such that

$$\lim_{N\to\infty} P(N)N^{\frac{1}{2}-\epsilon} > 1 \quad (134)$$

we have for $0 < \kappa < 1$

$$C_b^N \geq \frac{1}{4}\log\left(1 + \frac{\bar{h}_l^8}{4\bar{h}_u^4}(P(N))^2 N\right) \quad (135)$$

$$\geq \frac{\kappa}{4}\log\left((P(N))^2 N\right) \quad (136)$$

$$= \frac{\kappa}{4}\log(NP(N)) + \frac{\kappa}{4}\log(P(N)) \quad (137)$$

$$\geq \frac{\kappa}{4}\frac{4\epsilon}{1+2\epsilon}\log(NP(N)) \quad (138)$$

where the last step follows from

$$\frac{\kappa}{4}\left(1 - \frac{4\epsilon}{1+2\epsilon}\right)\log(NP(N)) + \frac{\kappa}{4}\log(P(N)) = \frac{\kappa}{4}\frac{2}{1+2\epsilon}\log(P(N)N^{\frac{1}{2}-\epsilon}) \geq 0 \quad (139)$$

when N is large enough, i.e., there exists an $N_4(\kappa) > 0$, such that when $N > N_4(\kappa)$, (135), (136), and (139) are true, and therefore, (138) is true.

Thus, combining all possible cases of $P(N)$, we see that when $P(N)$ is such that there exists a constant $\epsilon > 0$, such that

$$\lim_{N\to\infty} P(N)N^{\frac{1}{2}-\epsilon} > 1 \quad (140)$$

for any $0 < \kappa < 1$, the following rate C_a^N from sensor node i to the collector node is achievable:

$$C_a^N = \kappa\nu\log(NP(N)) \quad (141)$$

where constant ν is

$$\nu = \min\left(\frac{\epsilon}{1+2\epsilon}, \frac{1}{4}\right) \quad (142)$$

when N is large enough.

Since the achievable rate C_a^N is achievable for any sensor i, by a time sharing argument, the region

$$\sum_{i=1}^{N} R_i^N \leq C_a^N \quad (143)$$

is achievable.

For all other $P(N)$, from (124), we see that the achievable sum rate approaches a positive constant or zero as N goes to infinity.

D. Proof of Theorem 5.2

We restate the generalization of [16, Theorem 1], which appeared in [27, Theorem 1] for N sensor nodes below. This provides us with an achievable sum rate–distortion point, since the sum rate constraint is always tight [27].

Theorem A.1: [16], [17] A rate–distortion sum rate R_c and distortion D_c are achievable if there exist random variables $T_1, T_2, \ldots, T_N$ with

$$\left(s(u), u \in [0, U_0], S_{\{i\}^c}, T_{\{i\}^c}\right) \to S_i \to T_i, \quad i = 1, 2, \ldots, N \quad (144)$$

and an estimator function

$$\hat{S}(u) = g(T_1, T_2, \ldots, T_N) \quad (145)$$

such that

$$R_c \geq I(S_1, S_2, \ldots, S_N; T_1, T_2, \ldots, T_N) \quad (146)$$

$$D_c \geq E[d(S(u), g(T_1, T_2, \ldots, T_N))] \quad (147)$$

where random variables $\{S_i\}_{i=1}^N$ are defined as $S_i = S(u_i)$, $i = 1, 2, \ldots, N$.

We obtain an achievable rate–distortion point when we specify the relationship between $(S(u), \{S_i\}_{i=1}^\infty, \{T_i\}_{i=1}^\infty)$ as

$$T_i = S_i + W_i, \qquad i = 1, 2, \ldots, N \quad (148)$$

where W_i, $i = 1, 2, \ldots, N$, are i.i.d. Gaussian random variables with zero-mean and variance σ_D^2 and independent of everything else. Here, we can adjust σ_D^2 to achieve various feasible rate–distortion points [16].

We choose the minimum mean-square error (MMSE) estimator to estimate $S(u)$ from observations $\{T_k\}_{k=1}^N$. Hence, the achieved distortion is

$$D_c^N(\sigma_D^2) = \frac{1}{U_0}\int_0^{U_0}\left(K(u,u) - \boldsymbol{\rho}_N^T(u) \times \left(\boldsymbol{\Sigma}_N + \sigma_D^2\boldsymbol{I}\right)^{-1}\boldsymbol{\rho}_N(u)\right)du \quad (149)$$

The sum rate required to achieve this distortion is

$$R_c^N\left(\sigma_D^2\right) = I(S_1, S_2, \ldots, S_N; T_1, T_2, \ldots, T_N) \quad (150)$$

$$= \frac{1}{2}\log\det\left(I + \frac{1}{\sigma_D^2}\boldsymbol{\Sigma_N}\right)$$

$$= \sum_{k=0}^{N-1}\frac{1}{2}\log\left(1 + \frac{\mu_k^{(N)}}{\sigma_D^2}\right) \quad (151)$$

where $\mu_0^{(N)}, \mu_1^{(N)}, \ldots, \mu_{N-1}^{(N)}$ are the eigenvalues of $\boldsymbol{\Sigma_N}$.

Next, let $\theta = \frac{U_0}{N}\sigma_D^2$, $\boldsymbol{\Sigma}'_{\boldsymbol{N}} = \frac{U_0}{N}\boldsymbol{\Sigma}_{\boldsymbol{N}}$, and, therefore, $\mu_k^{(N)'} = \frac{U_0}{N}\mu_k^{(N)}$. We define two functions of θ as

$$R_a^N(\theta) \triangleq R_c\left(\sigma_D^2\right) = \sum_{k=0}^{N-1}\frac{1}{2}\log\left(1+\frac{\mu_k^{(N)'}}{\theta}\right) \tag{152}$$

and

$$\begin{aligned} D_a^N(\theta) &\triangleq D_c^N\left(\sigma_D^2\right) \\ &= \frac{1}{U_0}\int_0^{U_0}\Big(K(u,u) - \frac{U_0}{N}\boldsymbol{\rho}_N^T(u) \\ &\qquad \times \left(\boldsymbol{\Sigma}'_{\boldsymbol{N}} + \theta\boldsymbol{I}\right)^{-1}\boldsymbol{\rho}_N(u)\Big)\,du \end{aligned} \tag{153}$$

and by definition, sum rate $R_a^N(\theta)$ and distortion $D_a^N(\theta)$ are achievable for an arbitrary Gaussian random process.

E. Proof of Lemma 5.1

Using the matrix inversion lemma [18]

$$\left(\boldsymbol{\Sigma}'_{\boldsymbol{N}} + \theta\boldsymbol{I}\right)^{-1} = \boldsymbol{\Sigma}'^{-1}_{\boldsymbol{N}} - \boldsymbol{\Sigma}'^{-1}_{\boldsymbol{N}}\left(\frac{1}{\theta}\boldsymbol{I} + \boldsymbol{\Sigma}'^{-1}_{\boldsymbol{N}}\right)^{-1}\boldsymbol{\Sigma}'^{-1}_{\boldsymbol{N}} \tag{154}$$

we have

$$\begin{aligned} D_a^N(\theta) &= \frac{1}{U_0}\int_0^{U_0}\left(K(u,u) - \frac{U_0}{N}\boldsymbol{\rho}_N^T(u)\boldsymbol{\Sigma}'^{-1}_{\boldsymbol{N}}\boldsymbol{\rho}_N(u)\right)du \\ &\quad + \frac{1}{N}\int_0^{U_0}\boldsymbol{\rho}_N^T(u)\boldsymbol{\Sigma}'^{-1}_{\boldsymbol{N}}\left(\frac{1}{\theta}\boldsymbol{I} + \boldsymbol{\Sigma}'^{-1}_{\boldsymbol{N}}\right)^{-1} \\ &\quad \times \boldsymbol{\Sigma}'^{-1}_{N}\boldsymbol{\rho}_N(u)du \qquad (155) \\ &= D_s^{(N)} + D^{(N)}(\theta) \qquad (156) \end{aligned}$$

where we have defined $D_s^{(N)}$ and $D^{(N)}(\theta)$ as the first and second terms of the right-hand side of (155), respectively. We continue evaluating $D^{(N)}(\theta)$.

$$\begin{aligned} &D^{(N)}(\theta) \\ &= \frac{1}{N}\sum_{i=1}^{N}\int_{\frac{i-1}{N}U_0}^{\frac{i}{N}U_0}\left(\boldsymbol{\rho}_N^T\left(\frac{i-1}{N}U_0\right) - \boldsymbol{\Delta}_i^T(u)\right) \\ &\quad \times \boldsymbol{\Sigma}'^{-1}_{\boldsymbol{N}}\left(\frac{1}{\theta}\boldsymbol{I} + \boldsymbol{\Sigma}'^{-1}_{\boldsymbol{N}}\right)^{-1} \\ &\quad \times \boldsymbol{\Sigma}'^{-1}_{\boldsymbol{N}}\left(\boldsymbol{\rho}_N\left(\frac{i-1}{N}U_0\right) - \boldsymbol{\Delta}_i(u)\right)du \qquad (157) \\ &= \frac{1}{U_0}\sum_{i=1}^{N}\left(\left(\frac{1}{\theta}\boldsymbol{I} + \boldsymbol{\Sigma}'^{-1}_{\boldsymbol{N}}\right)^{-1}\right)_{(i,i)} \\ &\quad - 2\frac{1}{U_0}\sum_{i=1}^{N}\int_{\frac{i-1}{N}U_0}^{\frac{i}{N}U_0}\left(\left(\frac{1}{\theta}\boldsymbol{I} + \boldsymbol{\Sigma}'^{-1}_{\boldsymbol{N}}\right)^{-1}\boldsymbol{\Sigma}'^{-1}_{\boldsymbol{N}}\boldsymbol{\Delta}_i(u)\right)_i du \\ &\quad + \frac{1}{N}\sum_{i=1}^{N}\int_{\frac{i-1}{N}U_0}^{\frac{i}{N}U_0}\boldsymbol{\Delta}_i^T(u)\boldsymbol{\Sigma}'^{-1}_{\boldsymbol{N}}\left(\frac{1}{\theta}\boldsymbol{I} + \boldsymbol{\Sigma}'^{-1}_{\boldsymbol{N}}\right)^{-1} \\ &\quad \times \boldsymbol{\Sigma}'^{-1}_{\boldsymbol{N}}\boldsymbol{\Delta}_i(u)du \qquad (158) \end{aligned}$$

where $\Delta_i(u)$ is defined as

$$\boldsymbol{\Delta}_i(u) = \boldsymbol{\rho}_N\left(\frac{i-1}{N}U_0\right) - \boldsymbol{\rho}_N(u) \tag{159}$$

for $\frac{i-1}{N}U_0 \le u \le \frac{i}{N}U_0$, and (158) follows based on the fact that

$$\boldsymbol{\rho}_N^T\left(\frac{i-1}{N}U_0\right)\boldsymbol{\Sigma}'^{-1}_{\boldsymbol{N}} = \frac{N}{U_0}\boldsymbol{e}_i \tag{160}$$

where $\boldsymbol{e}_i$ is the row vector whose ith entry is 1 and all other entries are 0.

The eigenvalues of $\boldsymbol{\Sigma}'^{-1}_{\boldsymbol{N}}\left(\frac{1}{\theta}\boldsymbol{I} + \boldsymbol{\Sigma}'^{-1}_{\boldsymbol{N}}\right)^{-1}\boldsymbol{\Sigma}'^{-1}_{\boldsymbol{N}}$ are

$$\frac{\theta}{\mu_k^{(N)'} + \theta}\frac{1}{\mu_k^{(N)'}}, \qquad k = 0, 1, \ldots, N-1 \tag{161}$$

which are smaller than the corresponding eigenvalues of $\boldsymbol{\Sigma}'^{-1}_{\boldsymbol{N}}$, i.e., $\frac{1}{\mu_k^{(N)'}}$. Thus, the third term in (158) is bounded by

$$\begin{aligned} &\frac{1}{N}\sum_{i=1}^{N}\int_{\frac{i-1}{N}U_0}^{\frac{i}{N}U_0}\boldsymbol{\Delta}_i^T(u) \\ &\quad \times \boldsymbol{\Sigma}'^{-1}_{\boldsymbol{N}}\left(\frac{1}{\theta}\boldsymbol{I} + \boldsymbol{\Sigma}'^{-1}_{\boldsymbol{N}}\right)^{-1}\boldsymbol{\Sigma}'^{-1}_{\boldsymbol{N}}\boldsymbol{\Delta}_i(u)du \\ &\quad \le \frac{1}{N}\sum_{i=1}^{N}\int_{\frac{i-1}{N}U_0}^{\frac{i}{N}U_0}\boldsymbol{\Delta}_i^T(u)\boldsymbol{\Sigma}'^{-1}_{\boldsymbol{N}}\boldsymbol{\Delta}_i(u)du. \end{aligned} \tag{162}$$

To further upper bound the third term in (158), we write (163)–(166) at the top of the following page, where $A^{(N)}$ is defined in (44). Then, we have the third term in (158) upper bounded by $A^{(N)}$ because of (162), (166), and the fact that $D_s^{(N)}$ is nonnegative, i.e.,

$$\begin{aligned} &\frac{1}{N}\sum_{i=1}^{N}\int_{\frac{i-1}{N}U_0}^{\frac{i}{N}U_0}\boldsymbol{\Delta}_i^T(u)\boldsymbol{\Sigma}'^{-1}_{\boldsymbol{N}} \\ &\quad \times \left(\frac{1}{\theta}\boldsymbol{I} + \boldsymbol{\Sigma}'^{-1}_{\boldsymbol{N}}\right)^{-1}\boldsymbol{\Sigma}'^{-1}_{\boldsymbol{N}}\boldsymbol{\Delta}_i(u)du \le A^{(N)}. \end{aligned} \tag{167}$$

Furthermore, we can see from (166) that

$$D_s^{(N)} \le A^{(N)}. \tag{168}$$

Now, we evaluate the second term in (158). Since

$$\begin{aligned} &\left|\left(\left(\frac{1}{\theta}\boldsymbol{I} + \boldsymbol{\Sigma}'^{-1}_{\boldsymbol{N}}\right)^{-1}\boldsymbol{\Sigma}'^{-1}_{\boldsymbol{N}}\boldsymbol{\Delta}_i(u)\right)_i\right| \\ &\quad \le \left\|\left(\left(\frac{1}{\theta}\boldsymbol{I} + \boldsymbol{\Sigma}'^{-1}_{\boldsymbol{N}}\right)^{-1}\boldsymbol{\Sigma}'^{-1}_{\boldsymbol{N}}\boldsymbol{\Delta}_i(u)\right)_i\right\| \qquad (169) \\ &\quad \le \left\|\left(\frac{1}{\theta}\boldsymbol{I} + \boldsymbol{\Sigma}'^{-1}_{\boldsymbol{N}}\right)^{-1}\boldsymbol{\Sigma}'^{-1}_{\boldsymbol{N}}\right\|_2 \cdot \|\boldsymbol{\Delta}_i(u)\| \qquad (170) \\ &\quad = \max_{0\le k\le N-1}\left(\mu_k^{(N)'}\right)^{-1}\left(\frac{1}{\theta} + \frac{1}{\mu_k^{(N)'}}\right)^{-1}\|\boldsymbol{\Delta}_i(u)\| \qquad (171) \\ &\quad \le \|\boldsymbol{\Delta}_i(u)\|. \qquad (172) \end{aligned}$$

$$D_s^{(N)} = \frac{1}{U_0}\int_0^{U_0}\left(K(u,u) - \frac{U_0}{N}\boldsymbol{\rho}_N^T(u){\boldsymbol{\Sigma}'_{\boldsymbol{N}}}^{-1}\boldsymbol{\rho}_N(u)\right)du \tag{163}$$

$$= \frac{1}{U_0}\sum_{i=1}^{N}\int_{\frac{i-1}{N}U_0}^{\frac{i}{N}U_0}\left(K(u,u) - \frac{U_0}{N}\left(\boldsymbol{\rho}_N^T\left(\frac{i-1}{N}U_0\right) - \boldsymbol{\Delta}_i(u)^T\right){\boldsymbol{\Sigma}'_N}^{-1}\left(\boldsymbol{\rho}_N\left(\frac{i-1}{N}U_0\right) - \boldsymbol{\Delta}_i(u)\right)\right)du \tag{164}$$

$$= \frac{1}{U_0}\sum_{i=1}^{N}\int_{\frac{i-1}{N}U_0}^{\frac{i}{N}U_0}\left(K(u,u) - K\left(\frac{i-1}{N}U_0, \frac{i-1}{N}U_0\right)\right)du$$
$$+ \frac{2}{U_0}\sum_{i=1}^{N}\int_{\frac{i-1}{N}U_0}^{\frac{i}{N}U_0}\left(\boldsymbol{\Delta}_i(u)\right)_i du - \frac{1}{N}\sum_{i=1}^{N}\int_{\frac{i-1}{N}U_0}^{\frac{i}{N}U_0}\boldsymbol{\Delta}_i(u)^T{\boldsymbol{\Sigma}'_{\boldsymbol{N}}}^{-1}\boldsymbol{\Delta}_i(u)du \tag{165}$$

$$= A^{(N)} - \frac{1}{N}\sum_{i=1}^{N}\int_{\frac{i-1}{N}U_0}^{\frac{i}{N}U_0}\boldsymbol{\Delta}_i(u)^T{\boldsymbol{\Sigma}'_{\boldsymbol{N}}}^{-1}\boldsymbol{\Delta}_i(u)du \tag{166}$$

Therefore, the second term in (158) is bounded by

$$\left|-2\frac{1}{U_0}\sum_{i=1}^{N}\int_{\frac{i-1}{N}U_0}^{\frac{i}{N}U_0}\left(\left(\frac{1}{\theta}\boldsymbol{I} + {\boldsymbol{\Sigma}'_{\boldsymbol{N}}}^{-1}\right)^{-1} \times {\boldsymbol{\Sigma}'_{\boldsymbol{N}}}^{-1}\boldsymbol{\Delta}_i(u)\right)_i du\right| \le B^{(N)} \tag{173}$$

where $B^{(N)}$ is defined in (45). Finally, the first term in (158) can be written as

$$\frac{1}{U_0}\sum_{i=1}^{N}\left(\left(\frac{1}{\theta}\boldsymbol{I} + {\boldsymbol{\Sigma}'_{\boldsymbol{N}}}^{-1}\right)^{-1}\right)_{(i,i)} = \frac{1}{U_0}\sum_{k=0}^{N-1}\left(\frac{1}{\theta} + \frac{1}{\mu_k^{(N)'}}\right)^{-1} \triangleq D_b^N(\theta) \tag{174}$$

where the last step is by the definition of $D_b^N(\theta)$ in (46). Hence, for an arbitrary Gaussian random process, by (156), (158), (167), (168), (173), and (174), we have shown that

$$D_a^N(\theta) \le 2A^{(N)} + B^{(N)} + D_b^N(\theta) \tag{175}$$

F. Proof of Lemma 5.2

Using Condition 2 in Section III and the definition of $A^{(N)}$, $B^{(N)}$ in (44), (45), we have

$$A^{(N)} \le B\left(2^{\frac{\alpha}{2}} + 2\right)U_0^{\alpha}\frac{1}{N^{\alpha}} = \Theta\left(N^{-\alpha}\right) \tag{176}$$

$$B^{(N)} \le 2BU_0^{\alpha}\frac{N^{\frac{1}{2}}}{N^{\alpha}} = \Theta\left(N^{\frac{1}{2}-\alpha}\right) \tag{177}$$

G. Some Properties of $\mu_k^{(N)'}$

Lemma A.3: For all Gaussian random processes in $\mathcal{A}$, let $K_1(N)$ be a sequence of numbers that satisfies

$$\lim_{N\to\infty}\frac{1}{K_1(N)} = 0 \tag{178}$$

$$\lim_{N\to\infty}\frac{K_1(N)^{2\tau}}{N^{\gamma}} = 0 \tag{179}$$

$$\lim_{N\to\infty}\frac{K_1(N)^{x+1+\tau}}{N^{\beta}} = 0. \tag{180}$$

Then, for each k such that $k \le K_1(N)$, there exists an eigenvalue $\mu^{(N)'}$, different for each k, of $\boldsymbol{\Sigma}'_{\boldsymbol{N}}$ such that

$$\left|\mu^{(N)'} - \lambda_k\right| \le d_1\frac{(k+B_7)^{\tau}}{N^{\beta}} \tag{181}$$

for some $d_1 > 0$ and some positive integer B_7, both independent of k and N, when N is large enough.

Lemma A.3 shows that the convergence of $\mu_k^{(N)'}$ to λ_k is not uniform, and the approximation of $\mu_k^{(N)'}$ using λ_k is accurate only when $k \ll N^{\frac{\gamma}{2\tau}}$ and $\lambda_k \gg d_1\frac{(k+B_7)^{\tau}}{N^{\beta}}$. When the conditions of Lemma A.3 are satisfied, we label the $\mu^{(N)'}$ that satisfies (181) to be $\mu_k^{(N)'}$ for $k \le K_1(N)$. The remaining $N - K_1(N)$ eigenvalues of $\mu^{(N)'}$ will be labeled according to the order from large to small.

Lemma A.4: For all Gaussian random processes in $\mathcal{A}$, let two sequences ϑ_L^N and ϑ_U^N satisfy (51). Then, for any constant $0 < \kappa < 1$, we have

$$\sum_{k=\left\lfloor\left(\frac{d_u}{\theta}\right)^{\frac{1}{x}}+c_u\right\rfloor+1}^{N-1}\mu_k^{(N)'} \le \frac{d_u^{\frac{1}{x}}}{(x-1)\kappa^2}\theta^{1-\frac{1}{x}} \tag{182}$$

when $\theta \in \left[\vartheta_L^N, \vartheta_U^N\right]$ and N is large enough.

Lemma A.4 shows that the sum of the eigenvalues that do not converge to λ_k for $k = 0, 1, \ldots, \left\lfloor\left(\frac{d_u}{\theta}\right)^{\frac{1}{x}} + c_u\right\rfloor$ is the same order as

$$\sum_{k=\left\lfloor\left(\frac{d_u}{\theta}\right)^{\frac{1}{x}}+c_u\right\rfloor+1}^{\infty}\lambda''_k$$

as calculated in (75).

1) Proof of Lemma A.3: Due to the rather lengthy nature of this proof, here, we present the major steps and leave some of the details out. A complete proof can be found in [30].

By definition, λ_k for any k satisfies

$$\lambda_k\phi_k\left(\frac{l-1}{N}U_0\right) = \int_0^{U_0}K\left(\frac{l-1}{N}U_0, v\right)\phi_k(v)dv, \quad \forall\, l = 1, 2, \ldots, N. \tag{183}$$

We rewrite the right-hand-side of (183) by

$$\frac{U_0}{N}\sum_{i=1}^{N} K\left(\frac{l-1}{N}U_0, \frac{i-1}{N}U_0\right)\phi_k\left(\frac{i-1}{N}U_0\right)+\epsilon_N^k\left(\frac{l-1}{N}U_0\right), \quad \forall\, l=1,2,\ldots,N \tag{184}$$

where $\epsilon_N^k\left(\frac{l-1}{N}U_0\right)$ is defined as

$$\sum_{i=1}^{N}\int_{\frac{i-1}{N}U_0}^{\frac{i}{N}U_0}\left(K\left(\frac{l-1}{N}U_0, v\right)\phi_k(v) - K\left(\frac{l-1}{N}U_0, \frac{i-1}{N}U_0\right)\phi_k\left(\frac{i-1}{N}U_0\right)\right)dv. \tag{185}$$

Using (183) and (184), we have for any $l=1,2,\ldots,N$

$$\lambda_k\phi_k\left(\frac{l-1}{N}U_0\right) = \frac{U_0}{N}\sum_{i=1}^{N} K\left(\frac{l-1}{N}U_0, \frac{i-1}{N}U_0\right) \times\phi_k\left(\frac{i-1}{N}U_0\right)+\epsilon_N^k\left(\frac{l-1}{N}U_0\right). \tag{186}$$

Let us define vector $\boldsymbol{a}_k^{(N)}$ of length of N by defining its lth element to be $\sqrt{\frac{U_0}{N}}\epsilon_N^k\left(\frac{l-1}{N}U_0\right)$ and vector $\boldsymbol{b}_k^{(N)}$ of length of N by defining its lth element to be $\sqrt{\frac{U_0}{N}}\phi_k\left(\frac{l-1}{N}U_0\right)$, we have in matrix form

$$\lambda_k\boldsymbol{b}_k^{(N)} = \boldsymbol{\Sigma}_{\boldsymbol{N}}'\boldsymbol{b}_k^{(N)} + \boldsymbol{a}_k^{(N)}. \tag{187}$$

The links between the eigenvalues of $\boldsymbol{\Sigma}_{\boldsymbol{N}}'$ and the eigenvalues of $K(u,v)$, i.e., the λ_k's, will be determined using (187). To do this, we first bound three quantities $\left\|\boldsymbol{a}_k^{(N)}\right\|$, $\left\|\boldsymbol{b}_k^{(N)}\right\|$, $\left|\boldsymbol{b}_m^{(N)^T}\boldsymbol{b}_l^{(N)}\right|$ for $k,m,l\le K_1(N)$ and $m\ne l$.

Now, we analyze the norm of $\boldsymbol{a}_k^{(N)}$. From the definition of $\epsilon_N^k\left(\frac{l-1}{N}U_0\right)$ in (185), and the fact that the random process satisfies Condition 3 in Section III, we have

$$\left|\epsilon_N^k\left(\frac{l-1}{N}U_0\right)\right| \le B_2U_0^{1+\beta}\frac{(k+B_1)^\tau}{N^\beta}. \tag{188}$$

Thus, the norm of vector $\boldsymbol{a}_k^{(N)}$ is bounded by

$$\left\|\boldsymbol{a}_k^{(N)}\right\| \le B_2U_0^{3/2+\beta}\frac{(k+B_1)^\tau}{N^\beta}. \tag{189}$$

Now, we will calculate the norm of vector $\boldsymbol{b}_k^{(N)}$. We write

$$1=\int_0^{U_0}\phi_k^2(u)du = \sum_{i=1}^{N}\frac{U_0}{N}\phi_k^2\left(\frac{i-1}{N}U_0\right)+\delta_N^k \tag{190}$$

where δ_N^k is defined as

$$\delta_N^k = \sum_{i=1}^{N}\int_{\frac{i-1}{N}U_0}^{\frac{i}{N}U_0}\left(\phi_k^2(u)-\phi_k^2\left(\frac{i-1}{N}U_0\right)\right)du. \tag{191}$$

We first upper bound $|\phi_k(u)|$, for $u\in[0,U_0]$. Let $F_k(s)$ be defined as

$$\int_0^s \phi_k^2(u)du. \tag{192}$$

Then, by the mean value theorem on interval $[0,U_0]$, we have that there exists a $U'\in[0,U_0]$, such that

$$1=F_k(U_0)-F_k(0)=\phi_k^2(U'). \tag{193}$$

Hence, using Condition 3 in Section III, we have

$$|\phi_k(u)-\phi_k(U')|\le B_3(k+B_4)^\tau U_0^\gamma, \qquad u\in[0,U_0]. \tag{194}$$

Thus

$$\max_{u\in[0,U_0]}|\phi_k(u)|\le B_3(k+B_4)^\tau U_0^\gamma+1. \tag{195}$$

Using (16) and (195), we have for any $v_1,v_2\in[0,U_0]$

$$\left|\phi_k^2(v_1)-\phi_k^2(v_2)\right| \le 2\left(B_3(k+B_4)^\tau U_0^\gamma+1\right)B_3(k+B_4)^\tau\left|v_1-v_2\right|^\gamma. \tag{196}$$

The approximation error δ_N^k satisfies

$$|\delta_N^k|\le U_0^{1+\gamma}\frac{2\left(B_3(k+B_4)^\tau U_0^\gamma+1\right)B_3(k+B_4)^\tau}{N^\gamma} \tag{197}$$

$$\le U_0^{1+\gamma}\frac{2\left(B_3(K_1(N)+B_4)^\tau U_0^\gamma+1\right)B_3(K_1(N)+B_4)^\tau}{N^\gamma} \tag{198}$$

$$\le B_5 \tag{199}$$

where (197) follows from (196), and (199) is due to the fact that $K_1(N)$ satisfies (178) and (179), for a fixed constant B_5 that satisfies $0<B_5<1$, there exists an integer $N_0>0$, such that for $N\ge N_0$

$$U_0^{1+\gamma}\frac{2\left(B_3(K_1(N)+B_4)^\tau U_0^\gamma+1\right)B_3(K_1(N)+B_4)^\tau}{N^\gamma}\le B_5. \tag{200}$$

Finally, by the definition of $\boldsymbol{b}_k^{(N)}$, we have

$$\left\|\boldsymbol{b}_k^{(N)}\right\| = \sqrt{\sum_{i=1}^{N}\frac{U_0}{N}\phi_k^2\left(\frac{i-1}{N}U_0\right)} = \sqrt{1-\delta_N^k} \tag{201}$$

where (201) follows from (190). From (199), we have

$$\sqrt{1-B_5}\le\left\|\boldsymbol{b}_k^{(N)}\right\|\le\sqrt{1+B_5}. \tag{202}$$

Next, we show that based on the orthogonality of the eigenfunctions of $\phi_k(t)$, the sampled version $\boldsymbol{b}_k^{(N)}$'s are almost orthogonal. Using (16) and (195), we have

$$|\phi_m(v_1)\phi_l(v_1)-\phi_m(v_2)\phi_l(v_2)| \le \left(2B_3^2(m+B_4)^\tau(l+B_4)^\tau U_0^\gamma + B_3(l+B_4)^\tau+B_3(m+B_4)^\tau\right)|v_1-v_2|^\gamma. \tag{203}$$

Let m and l be two different integers, that belong to $\{1, 2, \ldots, N\}$. Then, we have

$$\begin{aligned} 0 &= \int_0^{U_0} \phi_m(u)\phi_l(u)du \\ &= \sum_{i=1}^{N} \frac{U_0}{N}\phi_m\left(\frac{i-1}{N}U_0\right)\phi_l\left(\frac{i-1}{N}U_0\right) + \varepsilon_N^{m,l}. \end{aligned} \tag{204}$$

Then, we have (205)–(207) at the bottom of the page, where (205) follows from (203), (207) follows when N is large enough due to the fact that $K_1(N)$ satisfies (178), i.e., there exists an integer N_2 such that when $N > N_2$, (207) is true. The right-hand side of (207) converges to zero as N goes to infinity due to the fact that $K_1(N)$ satisfies (179). We have

$$\begin{aligned} \left|\boldsymbol{b}_m^{(N)^T}\boldsymbol{b}_l^{(N)}\right| &= \left|\sum_{i=1}^{N} \frac{U_0}{N}\phi_m\left(\frac{i-1}{N}U_0\right)\phi_l\left(\frac{i-1}{N}U_0\right)\right| \\ &= \left|\varepsilon_N^{m,l}\right| \qquad (208) \\ &\le \frac{2B_3^2(K_1(N)+B_4)^{2\tau}U_0^{1+2\gamma}}{N^\gamma} \qquad (209) \end{aligned}$$

which means that vectors $\boldsymbol{b}_m^{(N)}$ and $\boldsymbol{b}_l^{(N)}$ become more orthogonal as N gets larger.

Now, we are ready to establish the link between the eigenvalues of $\boldsymbol{\Sigma}'_{\boldsymbol{N}}$ and λ_k. From (187), we have

$$\begin{aligned} \left\|\boldsymbol{b}_k^{(N)}\right\| &\le \left\|\left(\boldsymbol{\Sigma}'_{\boldsymbol{N}} - \lambda_k\boldsymbol{I}\right)^{-1}\right\|_2 \left\|\boldsymbol{a}_k^{(N)}\right\| \qquad (210) \\ &= \left(\min_{0\le m\le N-1}\left|\mu_m^{(N)'} - \lambda_k\right|\right)^{-1}\left\|\boldsymbol{a}_k^{(N)}\right\|. \qquad (211) \end{aligned}$$

Thus, we have

$$\begin{aligned} \min_{0\le m\le N-1}\left|\mu_m^{(N)'} - \lambda_k\right| &\le \frac{\left\|\boldsymbol{a}_k^{(N)}\right\|}{\left\|\boldsymbol{b}_k^{(N)}\right\|} \qquad (212) \\ &\le \frac{B_2U_0^{3/2+\beta}}{\sqrt{1-B_5}}\frac{(k+B_1)^\tau}{N^\beta} \qquad (213) \\ &= d_0\frac{(k+B_1)^\tau}{N^\beta} \qquad (214) \end{aligned}$$

where (214) follows by defining d_0 as

$$d_0 = \frac{B_2U_0^{3/2+\beta}}{\sqrt{1-B_5}}. \tag{215}$$

Hence, for $k = 0, 1, 2, \ldots, K_1(N)$, there exists an eigenvalue $\mu^{(N)'}$ of $\boldsymbol{\Sigma}'_{\boldsymbol{N}}$ such that

$$\left|\mu^{(N)'} - \lambda_k\right| \le d_0\frac{(k+B_1)^\tau}{N^\beta} \tag{216}$$

when N is large enough.

For $k = 0, 1, 2, \ldots, K_1(N)$, if we label the $\mu^{(N)'}$ that satisfies (216) to be $\mu_k^{(N)'}$, then when λ_k for different k's are sufficiently close; more specifically

$$|\lambda_m - \lambda_l| \le 2d_0\frac{(K_1(N)+B_1)^\tau}{N^\beta}, \quad m, l \le K_1(N), m \ne l \tag{217}$$

$\mu_m^{(N)'}$ and $\mu_l^{(N)'}$, though labeled differently, might be the same eigenvalue of $\boldsymbol{\Sigma}'_{\boldsymbol{N}}$, which is undesirable. If we relax the minimum distance of $d_0\frac{(k+B_1)^\tau}{N^\beta}$, we will be able to eliminate this problem. Thus, we will next show that for $k = 0, 1, 2, \ldots, K_1(N)$, there exists an eigenvalue $\mu^{(N)'}$ of $\boldsymbol{\Sigma}'_{\boldsymbol{N}}$, different for each k, such that

$$\left|\mu^{(N)'} - \lambda_k\right| \le (2\bar{\chi}+1)\sqrt{d_2}d_0\frac{(k+\bar{\chi}+B_1)^\tau}{N^\beta} \tag{218}$$

when N is large enough, where we define $\bar{\chi} \triangleq \max(K_0 + 1 + c_u + c_l, 2c_u + 2c_l + 1)$ and the constant d_2 as the largest root of the following second-order equation:

$$\begin{aligned} (1-B_5)d_2^2 - 2\left((1-B_5) + 3\bar{\chi}(1+B_5)\right)d_2 \\ +(1-B_5) + 2\bar{\chi}(1+B_5) = 0. \end{aligned} \tag{219}$$

It can be checked that both roots of the preceding equation are real, and the largest root is a positive constant, strictly larger than $\frac{2\bar{\chi}(1+B_5)}{1-B_5} + 1$, that is a function of $\bar{\chi}$ and B_5.

First, let us define a cluster of λ's. We say that χ λ's are a cluster, where with no loss of generality, we may label these λ's $\lambda_k, \lambda_{k+1}, \ldots, \lambda_{k+\chi-1}$, if

$$\begin{aligned} \lambda_{k+l} - \lambda_{k+l+1} \le 2\sqrt{d_2}d_0\frac{(k+\bar{\chi}+B_1)^\tau}{N^\beta}, \\ 0 \le k \le K_1(N),\ l = 0, 1, \ldots, \chi-1. \end{aligned} \tag{220}$$

Note here that whether the λ's are in a cluster depends on N. By using Condition 1 in Section III and the fact that $K_1(N)$ satisfies (178) and (180), it can be shown that for large enough N, the size of a cluster is at most $\bar{\chi}$, which is a finite number.

Let the eigenvalues and the corresponding eigenvectors of $\boldsymbol{\Sigma}'_{\boldsymbol{N}}$ be $\mu_i^{(N)'}$ and $\boldsymbol{u}_i^{(N)}$, $i = 0, 1, \ldots, N-1$, with arbitrary

$$\begin{aligned} \left|\varepsilon_N^{m,l}\right| &\le U_0^{1+\gamma}\frac{2B_3^2(m+B_4)^\tau(l+B_4)^\tau U_0^\gamma + B_3(l+B_4)^\tau + B_3(m+B_4)^\tau}{N^\gamma} \qquad (205) \\ &\le 2U_0^{1+\gamma}\frac{B_3^2(K_1(N)+B_4)^{2\tau}U_0^{1+\gamma} + B_3(K_1(N)+B_4)^\tau}{N^\gamma} \qquad (206) \\ &\le \frac{2B_3^2(K_1(N)+B_4)^{2\tau}U_0^{1+2\gamma}}{N^\gamma} \qquad (207) \end{aligned}$$

labeling of the eigenvalues and eigenvectors. Following from (187), we have

$$\sum_{i=0}^{N-1}\left(\lambda_k-\mu_i^{(N)'}\right)\boldsymbol{u}_i^{(N)}\boldsymbol{u}_i^{(N)^T}\boldsymbol{b}_k^{(N)}=\boldsymbol{a}_k^{(N)}. \tag{221}$$

We take the norm squared on both sides, and due to the orthogonality of eigenvectors $\boldsymbol{u}_i^{(N)}$, we have

$$\sum_{i=0}^{N-1}\left(\lambda_k-\mu_i^{(N)'}\right)^2\left(\boldsymbol{u}_i^{(N)^T}\boldsymbol{b}_k^{(N)}\right)^2=\left\|\boldsymbol{a}_k^{(N)}\right\|^2, \quad k=0,1,2,\ldots \tag{222}$$

and we also have

$$\sum_{i=0}^{N-1}\left(\boldsymbol{u}_i^{(N)^T}\boldsymbol{b}_k^{(N)}\right)^2=\left\|\boldsymbol{b}_k^{(N)}\right\|^2, \qquad k=0,1,2,\ldots. \tag{223}$$

Let $\lambda_k,\lambda_{k+1},\ldots,\lambda_{k+\chi-1}$ be a cluster, and from previous arguments, we know $\chi\le\bar{\chi}$. Furthermore, we are only interested in the first $K_1(N)+1$ eigenvalues, and therefore $k+\chi-1\le K_1(N)$. We will prove this by contradiction. Suppose that only ς number of $\mu_i^{(N)'}$'s are within distance

$$\sqrt{d_2}d_0\frac{(k+\bar{\chi}+B_1)^\tau}{N^\beta} \tag{224}$$

from any of the $\lambda_k,\lambda_{k+1},\ldots,\lambda_{k+\chi-1}$, with $1\le\varsigma<\chi$, we will show that there is a contradiction, and therefore, we can conclude that our assumption that $\varsigma<\chi$ number of $\mu_i^{(N)'}$'s are within distance (224) from any of the $\lambda_k,\lambda_{k+1},\ldots,\lambda_{k+\chi}$ is not correct.

Let us label the $\mu^{(N)'}$ that are within distance (224) from any of the $\lambda_k,\lambda_{k+1},\ldots,\lambda_{k+\chi}$ $\mu_0^{(N)'},\mu_1^{(N)'},\ldots,\mu_{\varsigma-1}^{(N)'}$. Before we dive into the details, let us first explain the basic idea of the proof. $\boldsymbol{u}_0^{(N)},\boldsymbol{u}_1^{(N)},\ldots,\boldsymbol{u}_{\varsigma-1}^{(N)}$ form the basis of a ς-dimensional subspace. On the other hand, $\boldsymbol{b}_k^{(N)},\boldsymbol{b}_{k+1}^{(N)},\ldots,\boldsymbol{b}_{k+\chi-1}^{(N)}$ are almost orthogonal, according to (209), and roughly form the basis of a χ-dimensional subspace. Since all other $\mu_i^{(N)'}$'s, for $i=\varsigma,\varsigma+1,\ldots,N-1$, are farther than distance (224) away, by Wedin's theorem in perturbation theory [31], the angle between $\boldsymbol{b}_{k+l}^{(N)}$ and the subspace is small, for all $l=0,1,\ldots,\chi-1$. But this is not possible, since ς is strictly smaller than χ. Now, we proceed with the rigorous proof.

Note that $\varsigma\ge1$ because we have already proved (216). Based on (214), the distance in (224) satisfies

$$\sqrt{d_2}d_0\frac{(k+\bar{\chi}+B_1)^\tau}{N^\beta}\ge\frac{\sqrt{d_2}\left\|\boldsymbol{a}_{k+l}^{(N)}\right\|}{\left\|\boldsymbol{b}_{k+l}^{(N)}\right\|}, \quad l=0,1,\ldots,\chi-1. \tag{225}$$

Then, based on (222), we have

$$\left(\frac{\sqrt{d_2}\left\|\boldsymbol{a}_{k+l}^{(N)}\right\|}{\left\|\boldsymbol{b}_{k+l}^{(N)}\right\|}\right)^2\sum_{i=\varsigma}^{N-1}\left(\boldsymbol{u}_i^{(N)^T}\boldsymbol{b}_{k+l}^{(N)}\right)^2\le\left\|\boldsymbol{a}_{k+l}^{(N)}\right\|^2, \quad l=0,1,\ldots,\chi-1. \tag{226}$$

Hence, we have

$$\sum_{i=\varsigma}^{N-1}\left(\boldsymbol{u}_i^{(N)^T}\boldsymbol{b}_{k+l}^{(N)}\right)^2\le\frac{\left\|\boldsymbol{b}_{k+l}^{(N)}\right\|^2}{d_2}, \quad l=0,1,\ldots,\chi-1. \tag{227}$$

Together with (223), we have

$$\sum_{i=0}^{\varsigma-1}\left(\boldsymbol{u}_i^{(N)^T}\boldsymbol{b}_{k+l}^{(N)}\right)^2\ge\frac{(d_2-1)\left\|\boldsymbol{b}_{k+l}^{(N)}\right\|^2}{d_2}, \quad l=0,1,\ldots,\chi-1. \tag{228}$$

Since the $\boldsymbol{u}_i^{(N)}$ form a complete set of orthonormal basis in $\mathbb{R}^N$, we can write $\boldsymbol{b}_{k+l}^{(N)}$ as

$$\boldsymbol{b}_{k+l}^{(N)}=\sum_{i=0}^{\varsigma-1}\alpha_{k+l,i}\boldsymbol{u}_i^{(N)}+\boldsymbol{v}_{k+l}^{(N)}, \quad l=0,1,\ldots,\chi-1 \tag{229}$$

where $\boldsymbol{v}_{k+l}^{(N)}$ is orthogonal to $\boldsymbol{u}_i^{(N)}$, for $i=1,2,\ldots,\varsigma$. From (228) and (229), we have

$$\sum_{i=0}^{\varsigma-1}(\alpha_{k+l,i})^2\ge\frac{(d_2-1)\left\|\boldsymbol{b}_{k+l}^{(N)}\right\|^2}{d_2}, \qquad l=,0,1,2,\ldots,\chi-1 \tag{230}$$

$$\left\|\boldsymbol{v}_{k+l}^{(N)}\right\|^2\le\frac{\left\|\boldsymbol{b}_{k+l}^{(N)}\right\|^2}{d_2}, \qquad l=0,1,\ldots,\varsigma-1. \tag{231}$$

Furthermore, by using (209), (202), (231), and the fact that $K_1(N)$ satisfies (179), for $m,l=0,1,\ldots,\varsigma-1,m\ne l$, we have

$$\left|\sum_{i=0}^{\varsigma-1}\alpha_{k+m,i}\alpha_{k+l,i}\right|\le\frac{2(1+B_5)}{d_2} \tag{232}$$

when N is large enough.

Let us define matrix $\boldsymbol{A}$ to be of dimension $\varsigma\times\varsigma$, with the (i,j)th component being $\alpha_{k+i-1,j}$, and define vectors $\boldsymbol{b}$, $\boldsymbol{v}$, and $\boldsymbol{u}$ to be column vectors of length ς, with the ith element being $\boldsymbol{b}_{k+i-1}^{(N)^T}\boldsymbol{b}_{k+\varsigma}^{(N)}$, $\boldsymbol{v}_{k+i-1}^{(N)^T}\boldsymbol{b}_{k+\varsigma}^{(N)}$, and $\boldsymbol{u}_{i-1}^{(N)^T}\boldsymbol{b}_{k+\varsigma}^{(N)}$, respectively. Then, by (229), we have

$$\boldsymbol{b}=\boldsymbol{A}\boldsymbol{u}+\boldsymbol{v} \tag{233}$$

thus, we have

$$||\boldsymbol{u}||^2\le\left\|\boldsymbol{A}^{-1}\right\|_2^2(||\boldsymbol{b}||+||\boldsymbol{v}||)^2. \tag{234}$$

We start by evaluating $||\boldsymbol{A}^{-1}||_2^2$, which is equal to the inverse of the smallest eigenvalue of $\boldsymbol{A}^T\boldsymbol{A}$. From the definition of matrix $\boldsymbol{A}$, we have

$$\boldsymbol{A}^T\boldsymbol{A}=\boldsymbol{D}+\boldsymbol{E} \tag{235}$$

where $\boldsymbol{D}$ is an $\varsigma\times\varsigma$ diagonal matrix with the lth diagonal element being $\sum_{i=0}^{\varsigma-1}(\alpha_{k+l-1,i})^2$, and $\boldsymbol{E}$ is an $\varsigma\times\varsigma$ matrix with zero diagonals and (m,l)th element being

$\sum_{i=0}^{\varsigma-1} \alpha_{k+m-1,i}\alpha_{k+l-1,i}$, when $m \neq l$. The absolute difference between the smallest eigenvalue of $\boldsymbol{A}^T\boldsymbol{A}$ and $\boldsymbol{D}$ is upper bounded by $||\boldsymbol{E}||_2$ [31]. The smallest eigenvalue of $\boldsymbol{D}$ is

$$\min_{l\in\{0,1,\ldots,\varsigma-1\}} \sum_{i=0}^{\varsigma-1} (\alpha_{k+l,i})^2 \geq \min_{l} \frac{(d_2-1)\left|\left|\boldsymbol{b}_{k+l}^{(N)}\right|\right|^2}{d_2} \quad (236)$$

$$\geq \frac{(d_2-1)(1-B_5)}{d_2} \quad (237)$$

where (236) follows from (230), and (237) follows from (202) when $N > N_0$ since $k+\varsigma-1 \leq K_1(N)$. We can upper bound the spectral norm of matrix $\boldsymbol{E}$, i.e., $||\boldsymbol{E}||_2$, by the Frobenius norm of $\boldsymbol{E}$, i.e.,

$$||\boldsymbol{E}||_2^2 \leq \sum_{m\neq l} \left(\sum_{i=0}^{\varsigma-1} \alpha_{k+m-1,i}\alpha_{k+l-1,i}\right)^2 \quad (238)$$

$$\leq \varsigma^2 \left(\frac{2(1+B_5)}{d_2}\right)^2 \quad (239)$$

$$< \chi^2 \frac{4(1+B_5)^2}{d_2^2} \quad (240)$$

$$\leq \bar{\chi}^2 \frac{4(1+B_5)^2}{d_2^2} \quad (241)$$

where (239) follows from (232). Hence, we may conclude that

$$\left|\left|\boldsymbol{A}^{-1}\right|\right|_2^2 < \left(\frac{(d_2-1)(1-B_5)}{d_2} - \frac{2\bar{\chi}(1+B_5)}{d_2}\right)^{-1} \quad (242)$$

where the right-hand side is a positive number, by the definition of d_2. Next, we evaluate $||\boldsymbol{v}||^2$.

$$||\boldsymbol{v}||^2 = \sum_{i=0}^{\varsigma-1} \left(\boldsymbol{v}_{k+i}^{(N)^T} \boldsymbol{b}_{k+\varsigma}^{(N)}\right)^2 \quad (243)$$

$$\leq \sum_{i=0}^{\varsigma-1} \left|\left|\boldsymbol{v}_{k+i}^{(N)^T}\right|\right|^2 \left|\left|\boldsymbol{b}_{k+\varsigma}^{(N)}\right|\right|^2 \quad (244)$$

$$\leq \frac{\left|\left|\boldsymbol{b}_{k+\varsigma}^{(N)}\right|\right|^2}{d_2} \sum_{i=0}^{\varsigma-1} \left|\left|\boldsymbol{b}_{k+i}^{(N)}\right|\right|^2 \quad (245)$$

$$\leq \frac{\varsigma(1+B_5)}{d_2} \left|\left|\boldsymbol{b}_{k+\varsigma}^{(N)}\right|\right|^2 \quad (246)$$

$$< \frac{\chi(1+B_5)}{d_2} \left|\left|\boldsymbol{b}_{k+\varsigma}^{(N)}\right|\right|^2 \quad (247)$$

$$\leq \frac{\bar{\chi}(1+B_5)}{d_2} \left|\left|\boldsymbol{b}_{k+\varsigma}^{(N)}\right|\right|^2 \quad (248)$$

where (245) follows from (231), and (246) follows from (202) when $N > N_0$ since $k+\varsigma-1 \leq K_1(N)$. Finally, we evaluate $||\boldsymbol{b}||^2$.

$$||\boldsymbol{b}||^2 = \sum_{i=0}^{\varsigma-1} \left(\boldsymbol{b}_{k+i}^{(N)^T} \boldsymbol{b}_{k+\varsigma}^{(N)}\right)^2 \quad (249)$$

$$\leq \varsigma \left(\frac{2B_3^2(K_1(N)+B_4)^{2\tau}U_0^{1+2\gamma}}{N^\gamma}\right)^2 \quad (250)$$

$$< \chi \left(\frac{2B_3^2(K_1(N)+B_4)^{2\tau}U_0^{1+2\gamma}}{N^\gamma}\right)^2 \quad (251)$$

$$\leq \bar{\chi} \left(\frac{2B_3^2(K_1(N)+B_4)^{2\tau}U_0^{1+2\gamma}}{N^\gamma}\right)^2 \quad (252)$$

where (250) follows from (209) when $N > N_2$.

Following from (234), using (248), (252) and (242), we have

$$||\boldsymbol{u}||^2 \leq \left|\left|\boldsymbol{A}^{-1}\right|\right|_2^2 (||\boldsymbol{b}|| + ||\boldsymbol{v}||)^2 \quad (253)$$

$$< \left(\frac{(d_2-1)(1-B_5)}{d_2} - \frac{2\bar{\chi}(1+B_5)}{d_2}\right)^{-1} \times \left(2\sqrt{\frac{\bar{\chi}(1+B_5)}{d_2}} \left|\left|\boldsymbol{b}_{k+\varsigma}^{(N)}\right|\right|\right)^2 \quad (254)$$

$$= \frac{d_2-1}{d_2} \left|\left|\boldsymbol{b}_{k+\varsigma}^{(N)}\right|\right|^2 \quad (255)$$

where (254) follows when N is large enough, due to the fact that $K_1(N)$ satisfies (179), i.e., there exists an integer N_5, such that when $N > N_5$

$$\sqrt{\bar{\chi}} \left(\frac{2B_3^2(K_1(N)+B_4)^{2\tau}U_0^{1+2\gamma}}{N^\gamma}\right) \leq \frac{\sqrt{\bar{\chi}(1+B_5)}}{d_2}\sqrt{1-B_5} \quad (256)$$

and (254) is true, and (255) follows from the definition of d_2 by (219). Hence, when N is large enough, we have a contradiction with (228). Therefore, we conclude that there must be at least χ eigenvalues of $\boldsymbol{\Sigma}_{\boldsymbol{N}}'$ within distance (224) away from any of the clustered λ's, furthermore, from the definition of a cluster in (220), there must be at least χ eigenvalues within distance

$$(2\chi+1)\sqrt{d_2}d_0\frac{(k+\bar{\chi}+B_1)^\tau}{N^\beta} \quad (257)$$

which is less than or equal to

$$(2\bar{\chi}+1)\sqrt{d_2}d_0\frac{(k+\bar{\chi}+B_1)^\tau}{N^\beta} \quad (258)$$

away from all of the clustered λ's. We pick χ eigenvalues of $\boldsymbol{\Sigma}_{\boldsymbol{N}}'$ which are within distance (258) and arbitrarily pair each clustered λ with one of the eigenvalues. These eigenvalues will not be paired with any other λ because all other clusters of λs are at least distance $2\sqrt{d_2}d_0\frac{(k+\bar{\chi}+B_1)^\tau}{N^\beta}$ apart from this cluster.

Finally, by letting

$$d_1 = (2\bar{\chi}+1)\sqrt{d_2}d_0, \quad B_7 = \bar{\chi}+B_1 \quad (259)$$

we have the desired results when N is large enough. Note that B_7 is a positive integer and d_1 is a positive constant, independent of k and N.

2) Proof of Lemma A.4: In the proof of Lemma A.4, we will need results from Lemma A.2 and A.3. Thus, we will first

prove that under the condition of Lemma A.4, the results of Lemma A.2 and A.3 apply. Since

$$\lim_{N\to\infty} \vartheta_U^N = 0 \tag{260}$$

for any $0 < \kappa < 1$, when N is large enough, $\theta < \vartheta_U^N$ is small enough, which means that the result of Lemma A.2 is valid. Now we show that the result of Lemma A.3 is also true. Let $K_1(N) = (\frac{d_u}{\vartheta_L^N})^{\frac{1}{x}} + c_u$. Because of

$$\lim_{N\to\infty} \frac{1}{\vartheta_L^N N^{\frac{x\gamma}{2\tau}}} = 0, \quad \lim_{N\to\infty} \frac{1}{\vartheta_L^N N^{\frac{x\beta}{x+1+\tau}}} = 0 \tag{261}$$

we have (179) and (180). Because of (260) and the fact that $\vartheta_L^N \le \vartheta_U^N$, we have (178).

Hence, for any $0 \le k \le \left\lfloor \left(\frac{d_u}{\theta}\right)^{\frac{1}{x}} + c_u \right\rfloor$, the result of Lemma A.3 applies because

$$k \le \left\lfloor \left(\frac{d_u}{\vartheta_L^N}\right)^{\frac{1}{x}} + c_u \right\rfloor \le K_1(N) \tag{262}$$

and N is large enough.

Now, we will use the result of Lemma A.2 and A.3 to prove Lemma A.4. From the properties of the Karhunen–Loeve expansion, we know that

$$\sum_{k=0}^{\infty} \lambda_k = \int_0^{U_0} K(u,u)du < \infty. \tag{263}$$

Thus, for any constant $0 < \kappa < 1$, we have

$$\sum_{k=0}^{\left\lfloor (\frac{d_u}{\theta})^{\frac{1}{x}} + c_u \right\rfloor} \lambda_k = \sum_{k=0}^{\infty} \lambda_k - \sum_{k=\left\lfloor (\frac{d_u}{\theta})^{\frac{1}{x}} + c_u \right\rfloor + 1}^{\infty} \lambda_k \tag{264}$$

$$\ge \int_0^{U_0} K(u,u)du - \sum_{k=\left\lfloor (\frac{d_u}{\theta})^{\frac{1}{x}} + c_u \right\rfloor + 1}^{\infty} \lambda_k'' \tag{265}$$

$$\ge \int_0^{U_0} K(u,u)du - \frac{d_u^{\frac{1}{x}}}{(x-1)\kappa}\theta^{1-\frac{1}{x}} \tag{266}$$

where we have used (75) in Lemma A.2 to obtain (266).

From the definition of matrix $\mathbf{\Sigma_N}$, we have

$$\sum_{k=0}^{N-1} \mu_k^{(N)'} = \frac{U_0}{N}\mathrm{tr}\left(\mathbf{\Sigma_N}\right) = \frac{U_0}{N}\sum_{i=1}^{N} K\left(\frac{i-1}{N}U_0, \frac{i-1}{N}U_0\right). \tag{267}$$

Thus

$$\sum_{k=\left\lfloor (\frac{d_u}{\theta})^{\frac{1}{x}} + c_u \right\rfloor + 1}^{N-1} \mu_k^{(N)'} = \sum_{k=0}^{N-1} \mu_k^{(N)'} - \sum_{k=0}^{\left\lfloor (\frac{d_u}{\theta})^{\frac{1}{x}} + c_u \right\rfloor} \mu_k^{(N)'} \tag{268}$$

$$\le \frac{U_0}{N}\sum_{i=1}^{N} K\left(\frac{i-1}{N}U_0, \frac{i-1}{N}U_0\right) - \sum_{k=0}^{\left\lfloor (\frac{d_u}{\theta})^{\frac{1}{x}} + c_u \right\rfloor} \left(\lambda_k - d_1\frac{(k+B_7)^{\tau}}{N^{\beta}}\right) \tag{269}$$

$$\le \frac{U_0}{N}\sum_{i=1}^{N} K\left(\frac{i-1}{N}U_0, \frac{i-1}{N}U_0\right) - \int_0^{U_0} K(u,u)du + \frac{d_u^{\frac{1}{x}}}{(x-1)\kappa}\theta^{1-\frac{1}{x}} + \frac{d_1}{N^{\beta}} \sum_{k=0}^{\left\lfloor (\frac{d_u}{\theta})^{\frac{1}{x}} + c_u \right\rfloor} (k+B_7)^{\tau} \tag{270}$$

$$\le \frac{BU_0^{1+\alpha}2^{\frac{\alpha}{2}}}{N^{\alpha}} + \frac{d_u^{\frac{1}{x}}}{(x-1)\kappa}\theta^{1-\frac{1}{x}} + \frac{d_1\left(\left\lfloor (\frac{d_u}{\theta})^{\frac{1}{x}} + c_u \right\rfloor + 1\right)\left(\left\lfloor (\frac{d_u}{\theta})^{\frac{1}{x}} + c_u \right\rfloor + B_7\right)^{\tau}}{N^{\beta}} \tag{271}$$

$$\le \frac{d_u^{\frac{1}{x}}}{(x-1)\kappa^2}\theta^{1-\frac{1}{x}} \tag{272}$$

where (269) follows by Lemma A.3. We have used (266) to obtain (270), and Condition 2 in Section III to obtain (271). Equation (272) holds when N is large enough, because $\theta \ge \vartheta_L^N$, and ϑ_L^N satisfies (51). Therefore, for any $0 < \kappa < 1$, (182) holds for $\theta \in \left[\vartheta_L^N, \vartheta_U^N\right]$ when N is large enough.

H. Proof of Lemma 5.3

Since the condition of Lemma 5.3 is the same as Lemma A.4, the results of Lemmas A.2–A.4 hold. By the same argument as Lemma A.2, Lemma A.1 holds as well.

We first prove (52). Since ϑ_L^N satisfies

$$\lim_{N\to\infty} \frac{1}{\vartheta_L^N N^{\frac{\beta x}{x+\tau+1}}} = 0 \Rightarrow \lim_{N\to\infty} \frac{1}{\vartheta_L^N N^{x}} = 0 \tag{273}$$

when N is large enough such that

$$\left\lfloor \left(\frac{d_u}{\theta}\right)^{\frac{1}{x}} + c_u \right\rfloor + 1 \le \left\lfloor \left(\frac{d_u}{\vartheta_L^N}\right)^{\frac{1}{x}} + c_u \right\rfloor + 1 < N - 1 \tag{274}$$

we can provide an upper bound on $R_a^N(\theta)$ by splitting the sum of N variables into two parts

$$R_a^N(\theta) = \sum_{k=0}^{\left\lfloor (\frac{d_u}{\theta})^{\frac{1}{x}} + c_u \right\rfloor} \frac{1}{2}\log\left(1 + \frac{\mu_k^{(N)'}}{\theta}\right) + \sum_{k=\left\lfloor (\frac{d_u}{\theta})^{\frac{1}{x}} + c_u \right\rfloor + 1}^{N-1} \frac{1}{2}\log\left(1 + \frac{\mu_k^{(N)'}}{\theta}\right). \tag{275}$$

For any $0 < \kappa < 1$, we start with the first term in (275). See (276)–(279) at the top of the following page, where (276) follows from Lemma A.3, (277) follows because the derivative of

$$\sum_{k=0}^{\left\lfloor \left(\frac{d_u}{\theta}\right)^{\frac{1}{x}}+c_u \right\rfloor} \frac{1}{2}\log\left(1+\frac{\mu_k^{(N)'}}{\theta}\right) \leq \sum_{k=0}^{\left\lfloor \left(\frac{d_u}{\theta}\right)^{\frac{1}{x}}+c_u \right\rfloor} \frac{1}{2}\log\left(1+\frac{\lambda_k}{\theta}+d_1\frac{(k+B_7)^\tau}{\theta N^\beta}\right) \tag{276}$$

$$\leq \sum_{k=0}^{\left\lfloor \left(\frac{d_u}{\theta}\right)^{\frac{1}{x}}+c_u \right\rfloor} \frac{1}{2}\log\left(1+\frac{\lambda_k''}{\theta}\right)+\frac{d_1}{2\theta N^\beta}\sum_{k=0}^{\left\lfloor \left(\frac{d_u}{\theta}\right)^{\frac{1}{x}}+c_u \right\rfloor}(k+B_7)^\tau \tag{277}$$

$$\leq \left(\frac{\log 2+x}{2\kappa}\right)d_u^{\frac{1}{x}}\theta^{-\frac{1}{x}}+\frac{d_1\left(\left\lfloor\left(\frac{d_u}{\theta}\right)^{\frac{1}{x}}+c_u\right\rfloor+1\right)\left(\left\lfloor\left(\frac{d_u}{\theta}\right)^{\frac{1}{x}}+c_u\right\rfloor+B_7\right)^\tau}{N^\beta} \tag{278}$$

$$\leq \left(\frac{\log 2+x}{2\kappa^2}\right)d_u^{\frac{1}{x}}\theta^{-\frac{1}{x}} \tag{279}$$

the function $\frac{1}{2}\log(1+x)$ is bounded by $1/2$ for $x \geq 0$, and the observation in (14), (278) follows because of (76) in Lemma A.2. Equation (279) holds when N is large enough, because $\theta \geq \vartheta_L^N$, and ϑ_L^N satisfies (51).

Now, we will study the second term of (275).

$$\sum_{k=\left\lfloor \left(\frac{d_u}{\theta}\right)^{\frac{1}{x}}+c_u \right\rfloor+1}^{N-1} \frac{1}{2}\log\left(1+\frac{\mu_k^{(N)'}}{\theta}\right) \leq \frac{1}{2}\sum_{k=\left\lfloor \left(\frac{d_u}{\theta}\right)^{\frac{1}{x}}+c_u \right\rfloor+1}^{N-1}\mu_k^{(N)'} \tag{280}$$

$$\leq \frac{d_u^{\frac{1}{x}}}{2(x-1)\kappa^2}\theta^{-\frac{1}{x}} \tag{281}$$

where in obtaining (280) and (281), we have used the fact that $\log(1+x) \leq x$ and (182) in Lemma A.4, respectively.

We combine the results of (279) and (281) and obtain

$$R_a^N(\theta) \leq \frac{d_u^{\frac{1}{x}}\left(x^2-(1-\log 2)x+(1-\log 2)\right)}{2(x-1)\kappa^2}\theta^{-\frac{1}{x}}. \tag{282}$$

Using similar methods, and Lemma A.1 in place of Lemma A.2, we may also lower bound $R_a^N(\theta)$ as

$$R_a^N(\theta) \geq \frac{\kappa x d_l^{\frac{1}{x}}}{4}\theta^{-\frac{1}{x}}. \tag{283}$$

For full details, please see [30].

Now we evaluate $D_b^N(\theta)$ for large enough N and $\theta \in \left[\vartheta_L^N, \vartheta_U^N\right]$, and prove (53).

$$D_b^N(\theta) = U_0^{-1}\sum_{k=0}^{N-1}\left(\frac{1}{\theta}+\frac{1}{\mu_k^{(N)'}}\right)^{-1} \tag{284}$$

$$= U_0^{-1}\sum_{k=0}^{k=\left\lfloor \left(\frac{d_u}{\theta}\right)^{\frac{1}{x}}+c_u \right\rfloor}\left(\frac{1}{\theta}+\frac{1}{\mu_k^{(N)'}}\right)^{-1} + U_0^{-1}\sum_{k=\left\lfloor \left(\frac{d_u}{\theta}\right)^{\frac{1}{x}}+c_u \right\rfloor+1}^{N-1}\left(\frac{1}{\theta}+\frac{1}{\mu_k^{(N)'}}\right)^{-1} \tag{285}$$

$$\leq U_0^{-1}\sum_{k=0}^{k=\left\lfloor \left(\frac{d_u}{\theta}\right)^{\frac{1}{x}}+c_u \right\rfloor}\theta + U_0^{-1}\sum_{k=\left\lfloor \left(\frac{d_u}{\theta}\right)^{\frac{1}{x}}+c_u \right\rfloor+1}^{N-1}\mu_k^{(N)'} \tag{286}$$

$$\leq U_0^{-1}\theta\left(\left\lfloor\left(\frac{d_u}{\theta}\right)^{\frac{1}{x}}+c_u\right\rfloor+1\right) + U_0^{-1}\frac{d_u^{\frac{1}{x}}}{(x-1)\kappa^2}\theta^{1-\frac{1}{x}} \tag{287}$$

$$\leq \frac{d_u^{\frac{1}{x}}\left(1+\kappa^2(x-1)\right)}{\kappa^3(x-1)U_0}\theta^{1-\frac{1}{x}} \tag{288}$$

where (285) follows because of the same reason as (275), and (286) follows because of the fact that for $a, b \geq 0$, $\left(\frac{1}{a}+\frac{1}{b}\right)^{-1} \leq \min(a,b)$, and (287) follows from (182) of Lemma A.4, (288) follows because $\theta < \vartheta_U^N$, and ϑ_U^N goes to zero as N goes to infinity. Therefore, for any $0 < \kappa < 1$, (52) and (53) are true for $\theta \in \left[\vartheta_L^N, \vartheta_U^N\right]$ when N is large enough.

I. Proof of Theorem 5.3

Note that (52) implies that

$$\frac{\kappa^x x^x d_l}{4^x R^x} \leq \theta_a^N(R) \leq \left(\frac{d_u^{\frac{1}{x}}\left(x^2-(1-\log 2)x+(1-\log 2)\right)}{2(x-1)\kappa^2}\right)^x R^{-x} \tag{289}$$

for large enough N and R in the interval of

$$\left[\frac{d_u^{\frac{1}{x}}\left(x^2-(1-\log 2)x+(1-\log 2)\right)}{2(x-1)\kappa^2}\left(\vartheta_U^N\right)^{-\frac{1}{x}}, \frac{\kappa x d_l^{\frac{1}{x}}}{4}\left(\vartheta_L^N\right)^{-\frac{1}{x}}\right]. \tag{290}$$

From the definition of $D_a(R)$ in (47), we have (291)–(294) at the top of the following page, where (292) follows from (43), (293) follows because of (53), (294) follows from (49), (50),

$$D_a(R) = D_a^N(\theta_a^N(R)) \tag{291}$$
$$\leq 2A^{(N)} + B^{(N)} + D_b^N(\theta_a^N(R)) \tag{292}$$
$$\leq 2A^{(N)} + B^{(N)} + \frac{d_u^{\frac{1}{x}}\left(1+\kappa^2(x-1)\right)}{\kappa^3(x-1)U_0}\left(\theta_a^N(R)\right)^{1-\frac{1}{x}} \tag{293}$$
$$\leq O\left(N^{-\alpha}\right) + O\left(N^{1/2-\alpha}\right) + \frac{d_u(1+\kappa^2(x-1))\left(x^2-(1-\log 2)x+(1-\log 2)\right)^{x-1}}{U_0\kappa^{2x+1}2^{x-1}(x-1)^x}R^{1-x} \tag{294}$$

(289), and the fact that R in (56) implies that R is in (290), and when R is in (290), $\theta_a^N(R)$ is in $\left[\vartheta_L^N, \vartheta_U^N\right]$. When R is in (56), we have that the third term in (294) is much larger than the sum of the first and second terms when N is large enough due to the fact that

$$\lim_{N\to\infty}\frac{1}{\vartheta_{LL}^N N^{\frac{(\alpha-1/2)x}{x-1}}} = 0 \Rightarrow \lim_{N\to\infty}\frac{1}{\vartheta_{LL}^N N^{\frac{\alpha x}{x-1}}} = 0. \tag{295}$$

Therefore, for $0 < \kappa < 1$, (55) is true for R in the interval of (56) when N is large enough.

J. Proof of Theorem 5.4

Pick the sequences ϑ_{LL}^N and ϑ_U^N as

$$\vartheta_{LL}^N = \left(\frac{\nu}{\frac{x d_l^{\frac{1}{x}}}{8}}\log NP(N)\right)^{-x}$$
$$\vartheta_U^N = \left(\frac{\nu}{\frac{d_u^{\frac{1}{x}}\left(x^2-(1-\log 2)x+(1-\log 2)\right)}{2(x-1)\kappa^2}}\log NP(N)\right)^{-x/2}. \tag{296}$$

Then, because $P(N)$ satisfies (37) and (57), ϑ_{LL}^N satisfies (54), and ϑ_U^N satisfies (51). According to (38), we have the achievable rate C_a^N in the interval of (56), and thus, when N is large enough, Theorem 5.3 applies. Therefore, when $P(N)$ satisfies (37) and (57), for any $0 < \kappa < 1$, (59) holds when N is large enough.

References

[1] N. Liu and S. Ulukus, "Optimal distortion-power tradeoffs in sensor networks: Gauss-Markov random processes," in *Proc. IEEE Int. Conf. Communications*, Istanbul, Turkey, Jun. 2006, pp. 1543–1548.

[2] N. Liu and S. Ulukus, "Optimal distortion-power tradeoffs in Gaussian sensor networks," in *Proc. IEEE Int. Symp. Information Theory*, Seattle, WA, Jul. 2006, pp. 1534–1538.

[3] P. Gupta and P. R. Kumar, "The capacity of wireless networks," *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 388–404, Mar. 2000.

[4] A. Scaglione and S. D. Servetto, "On the interdependence of routing and data compression in multi-hop sensor networks," *Wireless Netw.*, vol. 11, no. 1-2, pp. 149–160, 2005.

[5] D. Marco, M. Liu, E. J. Duarte-Melo, 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. 2nd Int. Workshop on Information Processing in Sensor Networks, IPSN'03*, Palo Alto, CA, Apr. 2003.

[6] D. Marco and D. L. Neuhoff, "Entropy of quantized data at high sampling rates," in *Proc. IEEE Int. Symp. Information Theory*, Adelaide, Australia, Sep. 2005, pp. 342–346.

[7] H. El Gamal, "On the scaling laws of dense wireless sensor networks: the data gathering channel," *IEEE Trans. Inf. Theory*, vol. 51, no. 3, pp. 1229–1234, Mar. 2005.

[8] A. Kashyap, L. A. Lastras-Montano, C. Xia, and Z. Liu, "Distributed source coding in dense sensor networks," in *Proc. Data Compression Conf. (DCC)*, Snowbird, UT, Apr. 2005, pp. 13–22.

[9] D. L. Neuhoff and S. S. Pradhan, "An upper bound to the rate of ideal distributed lossy source coding of densely sampled data," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, Dallas, TX, Sep. 2006, vol. 5, pp. 14–19.

[10] U. Grenander and G. Szego, *Toeplitz Forms and Their Applications*. Berkeley and Los Angeles: Univ. California Press, 1958.

[11] M. Gastpar and M. Vetterli, "Power, spatio-temporal bandwidth, and distortion in large sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 4, pp. 745–754, Apr. 2005.

[12] T. M. Cover, A. El Gamal, and M. Salehi, "Multiple access channels with arbitrarily correlated sources," *IEEE Trans. Inf. Theory*, vol. IT-26, no. 6, pp. 648–657, Nov. 1980.

[13] G. Dueck, "A note on the multiple access channels with correlated sources," *IEEE Trans. Inf. Theory*, vol. IT-27, no. 2, pp. 232–235, Mar. 1981.

[14] W. Kang and S. Ulukus, "A single-letter upper bound for the sum rate of multiple access channels with correlated sources," in *Proc. 39th Asilomar Conf. Signals, Systems and Computers*, Pacific Grove, CA, Oct. 2005, pp. 1732–1736.

[15] W. Kang and S. Ulukus, "An outer bound for multiple access channels with correlated sources," in *Proc. 40th Annu. Conf. Information Sciences and Systems*, Princeton, NJ, Mar. 2006, pp. 240–244.

[16] T. J. Flynn and R. M. Gray, "Encoding of correlated observations," *IEEE Trans. Inf. Theory*, vol. IT-33, no. 6, pp. 773–787, Nov. 1987.

[17] M. Gastpar and M. Vetterli, "On the capacity of large Gaussian relay networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 3, pp. 765–779, Mar. 2005.

[18] R. A. Horn and C. R. Johnson, *Matrix analysis*. Cambridge, U.K.: Cambridge University Press, 1985.

[19] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*. New York: McGraw-Hill, 1991.

[20] G. E. Uhlenbeck and L. S. Ornstein, "On the theory of Brownian motion," *Phys. Rev.*, vol. 36, pp. 823–841, 1930.

[21] M. C. Wang and G. E. Uhlenbeck, "On the theory of Brownian motion II," *Rev. Mod. Phys.*, vol. 17, pp. 323–342, 1945.

[22] J. L. Doob, "The Brownian movement and stochastic equations," *Ann. Math.*, vol. 43, no. 2, pp. 351–369, Apr. 1942.

[23] L. Breiman, *Probability*. Reading, MA: Addison-Wesley, 1968.

[24] I. Karatzas and S. E. Shreve, *Brownian Motion and Stochastic Calculus*. New York: Springer-Verlag, 1988.

[25] T. Berger, *Rate Distortion Theory*. Englewood Cliffs, NJ: Prentice Hall, 1971.

[26] İ. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Europ. Trans. Telecommun.*, vol. 10, pp. 585–595, Nov. 1999.

[27] J. Chen, X. Zhang, T. Berger, and S. B. Wicker, "An upper bound on the sum-rate distortion function and its corresponding rate allocation schemes for the CEO problem," *IEEE J. Sel. Areas Commun.*, vol. 22, no. 6, pp. 977–987, Aug. 2004.

[28] T. Berger, "Multi-terminal source coding," in *Information Theory Approach to Communications*, G. Longo, Ed. New York: Springer-Verlag, 1978, pp. 171–231.

[29] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley-Interscience, 1991.

[30] N. Liu, "Capacity Results for Wireless Networks: Effects of Correlation Cooperation, and Interference," Ph.D. dissertation, Univ. Maryland, College Park, 2007.

[31] G. W. Stewart, "Perturbation theory for the singular value decomposition," in *SVD and Signal Processing, II: Algorithms, Analysis and Applications*, R. Vaccaro, Ed. Amsterdam, The Netherlands: Elsevier Science, 1991, pp. 99–109.