

Ordered Transmissions Schemes for Detection in Spatially Correlated Wireless Sensor Networks

Sayan Sen Gupta¹, *Graduate Student Member, IEEE*, and Neelesh B. Mehta², *Fellow, IEEE*

Abstract—The ordered transmissions scheme requires fewer sensor nodes to transmit their measurements than the conventional unordered transmissions scheme (UTS) in which all nodes transmit. Yet, it achieves the same error probability as UTS. For the practically relevant scenario in which the measurements of the sensor nodes are spatially correlated, we present a novel correlation-aware ordered transmissions scheme (CA-OTS) for the binary hypothesis testing problem with Gaussian statistics. It uses the timer scheme to make the nodes transmit their measurements in the decreasing order of the absolute values of the measurements without any node knowing the measurements of other nodes. CA-OTS applies to the general case where the hypotheses differ in the mean vector and covariance matrix, and markedly reduces the number of transmissions. It differs from the literature that assumes that the measurements of the nodes, when conditioned on the hypotheses, are statistically independent or the covariance matrix has a special structure. When the mean vector or covariance matrix is the same for the two hypotheses, we propose novel refinements that require even fewer transmissions. We also derive insightful upper bounds for them that apply to a general product-correlation model.

Index Terms—Wireless sensor networks, detection, ordered transmissions, correlation, energy-efficiency.

I. INTRODUCTION

ENERGY-EFFICIENCY is a critical issue in the design of wireless sensor networks (WSNs). It ensures that these networks have the longevity required by their many compelling and diverse applications such as transportation and logistics, environmental monitoring, military surveillance, and healthcare [2]. Several techniques, such as censoring, sensor selection, on-off keying, duty cycling, and clustering, have been proposed in the literature to address this challenge [3]–[5]. They improve the energy-efficiency by curtailing the number of transmissions by the nodes. However, this invariably degrades performance.

A noteworthy exception is the ordered transmissions scheme (OTS), which reduces the average number of transmissions

by 50% compared to the error rate-optimal unordered transmissions scheme (UTS), in which all the sensor nodes transmit. It does so without any increase in the error probability [6]–[16]. OTS was first proposed in [6] for the binary hypothesis testing problem. It exploits the fact that the decision statistic at the fusion node (FN) can be written as the sum of the log-likelihood ratios (LLRs) of the measurements at the individual sensor nodes. In OTS, the nodes transmit in the decreasing order of the absolute values of their LLRs. This is accomplished by the timer-based multiple access scheme [17] in which each node sets a timer that is a monotone non-increasing function of a local, non-negative real number called metric. In OTS, the metric is the absolute value of the LLR. When a node's timer expires, it transmits its LLR to the FN. This ensures that the nodes transmit their LLRs in the decreasing order of their metrics in a distributed manner without any node knowing the metric of any other node. Every time the FN receives an LLR from a node, it decides on a hypothesis or waits for the next transmission. Once the FN decides, it broadcasts a control signal to stop the other nodes from transmitting and draining their battery energies.

A. Literature Survey on OTS

In [7], OTS is used for target detection in radar networks. In [8], only a sensor node whose absolute value of the LLR exceeds a threshold can transmit. Among such nodes, the one with the largest channel power gain transmits. This is repeated over multiple coherence intervals and the transmitting node's power is adapted to minimize the total energy consumed. In [9], just one observation is shown to be sufficient for OTS to decide on a hypothesis when the number of nodes is asymptotically large. In [10], an ordered transmissions-based sequential detection scheme for spectrum sensing is used to maximize the weighted sum throughput of the primary and secondary users. In [11], a linear regression problem involving multiple sensors is studied. The sensors transmit in the increasing order of the difference between their training data and observed data, and the goal is to minimize the mean squared error. In [12], a constraint is imposed on the number of transmissions to the FN. A combination of OTS and slotted Aloha for detecting a shift in the mean is studied in [13]. In [14], OTS is adapted for energy harvesting sensor networks and the problem of missing transmissions is addressed.

The above articles [7]–[14] assume that the measurements of the sensor nodes, when conditioned on the hypotheses, are mutually independent. However, in many practical

Manuscript received June 7, 2020; revised October 19, 2020 and December 13, 2020; accepted December 14, 2020. Date of publication December 24, 2020; date of current version March 17, 2021. This work was partially supported by a research grant from Intel Corp., Bangalore and by the Department of Science and Technology (DST)-Swaranajayanti Fellowship under Award DST/SJF/ETA-01/2014-15. This article was presented in part at the IEEE International Conference on Communications, 2020 [1]. The associate editor coordinating the review of this article and approving it for publication was L. Lampe. (Corresponding author: Sayan Sen Gupta.)

The authors are with the Department of Electrical Communication Engineering, Indian Institute of Science, Bengaluru 560012, India (e-mail: sayan.riju88@gmail.com; neeleshbmehta@gmail.com).

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/TCOMM.2020.3047087>.

Digital Object Identifier 10.1109/TCOMM.2020.3047087

deployments, the measurements can be spatially dependent or correlated [18]. For example, as WSNs become more ubiquitous and the cost of a sensor node decreases, these nodes are likely to be densely deployed [19]. In these cases, the decision statistic at the FN is no longer the sum of the individual node LLRs. Instead, it also consists of cross-terms involving products of the measurements of multiple nodes. Thus, ordering on the basis of the absolute values of the LLRs no longer works.

OTS for correlated measurements has received less attention in the literature. It has recently been studied in [15], [16] for a decomposable Gaussian graphical model. In this model, each node is a vertex. The measurements of any pair of non-adjacent vertices are independent when conditioned on the measurements of the other vertices. Furthermore, the graph can be decomposed into maximal cliques, which might have some vertices in common. In [15], [16], the problems of detecting a shift in the mean vector and a shift in the covariance matrix are separately studied. Both works exploit an elegant property of the graphs to group the vertices belonging to maximal cliques into clusters. All the nodes within a cluster transmit their measurements to the cluster head. Ordering is used to reduce the number of cluster heads that transmit.

However, for several commonly studied correlation models, the graph is complete, i.e., it is a maximal clique. In this case, all the nodes end up transmitting. One example is the uniform correlation model [20], [21]. In it, the correlation between any two nodes is the same. This model is a special case of the more general product-correlation model [22] in which the correlation between two nodes i and j can be written as $\theta_i\theta_j$, for $i \neq j$, $1 \leq i, j \leq N$, where $\theta_1, \theta_2, \dots, \theta_N \in (-1, 1)$ are arbitrary constants and N is the number of nodes. This happens, for example, when the measurements at the sensor nodes are linearly attenuated, noise corrupted versions of a source signal [23]–[25]. A different example is the exponential correlation model in which the correlation-coefficient between nodes i and j is $\rho^{|i-j|}$, for some $0 < \rho < 1$ [26]. The graphical model for it is a line. Every maximal clique consists of only two adjacent nodes, one of which is the cluster head. The other node transmits to its two adjacent cluster heads. Thus, the total number of transmissions is not reduced.

B. Focus and Contributions

We propose a novel correlation-aware OTS (CA-OTS) for the binary hypothesis testing framework. It applies to Gaussian statistics with a general correlation model in which the two hypotheses differ in their mean vectors and their covariance matrices. In CA-OTS, the metric of a node is the absolute value of its measurement. And, when the timer expires, a node transmits its measurement to the FN. We derive novel decision rules at the FN that reduce the average number of transmissions while achieving the same optimal error probability as conventional UTS.

Since the measurements are correlated, this choice of the metric can seem counter-intuitive at first since the absolute value of the measurement no longer fully captures how informative it is. A holistic understanding of OTS, which we put

forth below, is needed to understand this. OTS consists of three key components that need to be jointly designed. The first component is the metric, which determines the order in which the nodes should transmit. The second component is the payload of the packet that a node transmits to the FN when its timer expires. The third component is the decision rules at the FN, which enable it to make a decision early but using only the information that it has received thus far from an ordered subset of nodes.

We then present refined versions of CA-OTS for the shift-in-covariance and shift-in-mean binary hypothesis testing problems. These refinements bring out the importance of designing the metric. Their corresponding novel decision rules exploit the additional structure in the problem to improve performance or reduce complexity. In the shift-in-covariance problem, the covariance matrices for the two hypotheses are different while the mean vectors are the same. For the product-correlation model, we derive an insightful upper bound on the average number of transmissions. It shows that the average number of transmissions decreases to the lowest possible value of one as the signal-to-noise ratio (SNR), which we formally define later, increases. To the best of our knowledge, similar results are available in the literature only for independent measurements [7], [9].

In the shift-in-mean problem, the mean vectors for the two hypotheses are different while the covariance matrices are the same. For it as well, we derive an insightful upper bound on the average number of transmissions for the product-correlation model. It shows that the average number of transmissions decreases by at least 50% as the separation between the two mean vectors increases. While such results are available in the literature, they only apply to independent measurements [6] or the decomposable Gaussian graphical model [15].

Our benchmarking results show that CA-OTS and its refinements markedly reduce the average number of transmissions compared to UTS. This translates into an improvement in the energy-efficiency of the WSN – without compromising on its error probability. The results also bring out the impact of correlation among the measurements on the performance of all the schemes.

C. General Mathematical Principle and Comments

Separability of the decision statistic in terms of functions of the local measurements is the key mathematical principle behind OTS when the measurements are independent when conditioned on the hypotheses. This leads to a natural ordering of the sensor nodes in terms of the absolute values of the LLRs. Even for correlated measurements, this principle is exploited in [15] and [16], where the decision statistic is written in a form that is separable across cluster heads for the decomposable Gaussian graphical model. Our work shows that even if the decision statistic is not separable, as is the case for several correlation models [20], [23]–[25], an ordering can be imposed so long as the lower and upper bounds on the decision statistic are separable. This principle differs from model order reduction techniques such as reduced rank detection [27],

which assume that the FN has access to all the measurements and employ a reduced rank, lower complexity detector that maximizes the divergence between the hypotheses.

There are two key differences between our approach and those in [15], [16]. First, our approach applies to any general correlation model. Second, the ordering happens across all the sensor nodes. On the other hand, in [15], [16], all nodes within a cluster transmit to the cluster head and ordering happens only across the cluster heads. CA-OTS also differs from sequential detection in which the nodes transmit one by one, but in a random order [28].

D. Organization and Notations

Section II presents the system model. Section III specifies CA-OTS and its refinements. Simulation results are presented in Section IV, and are followed by our conclusions in Section V.

Notations: The probability of an event B is denoted by $\Pr(B)$, and the conditional probability of B given event D by $\Pr(B|D)$. The expectation with respect to a random variable (RV) X is denoted by $\mathbb{E}_X[\cdot]$. Matrices and vectors are denoted using boldface characters. The notation $\mathbf{y} \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{C})$ means that \mathbf{y} is a Gaussian random vector with mean $\boldsymbol{\mu}$ and a covariance matrix \mathbf{C} . For a matrix \mathbf{P} , P_{ij} denotes the entry in the i^{th} row and j^{th} column, $\det(\mathbf{P})$ denotes its determinant, and $\lambda_{\min}(\mathbf{P})$ and $\lambda_{\max}(\mathbf{P})$ denote its smallest and largest eigenvalues, respectively. For a vector \mathbf{v} , \mathbf{v}^T denotes its transpose and v_i denotes its i^{th} element. The identity matrix of size $N \times N$ is denoted by \mathbf{I}_N and the all-zero vector is denoted by $\mathbf{0}$. For a real number x , $\text{sgn}(x)$ denotes its sign, $\lceil x \rceil$ denotes the ceiling function, and $\lfloor x \rfloor$ denotes the floor function. For N continuous RVs, X_1, X_2, \dots, X_N , we use the order statistics notation to denote the r^{th} largest value by $X_{[r]}$ and the index of this RV by $[r]$. Therefore, $X_{[1]} > X_{[2]} > \dots > X_{[N]}$.

II. SYSTEM MODEL

Consider a WSN that consists of N sensor nodes and an FN. Let $\mathbb{N} = \{1, 2, \dots, N\}$ denote the set of indices of the sensor nodes. Time is divided into measurement rounds. In each round, a decision needs to be made by the FN on the basis of the measurements it receives from some or all of the sensor nodes.

We consider the binary hypothesis testing framework, which is a fundamental problem in signal detection theory [6], [9], [10]. In general, the signal models under the two hypotheses for Gaussian statistics are given by [15], [16]

$$\begin{aligned} H_0 : \mathbf{y} &\sim \mathcal{N}(\boldsymbol{\mu}^{(0)}, \mathbf{R}^{(0)}), \\ H_1 : \mathbf{y} &\sim \mathcal{N}(\boldsymbol{\mu}^{(1)}, \mathbf{R}^{(1)}), \end{aligned} \quad (1)$$

where $\mathbf{y} = [y_1, y_2, \dots, y_N]^T$, y_i is the measurement at node i , $\boldsymbol{\mu}^{(0)}$ and $\mathbf{R}^{(0)}$ are the mean vector and covariance matrix, respectively, under hypothesis H_0 , and $\boldsymbol{\mu}^{(1)}$ and $\mathbf{R}^{(1)}$ are the mean vector and covariance matrix, respectively, under hypothesis H_1 . Let $\boldsymbol{\mu}^{(h)} = [\mu_1^{(h)}, \mu_2^{(h)}, \dots, \mu_N^{(h)}]^T$, for $h \in \{0, 1\}$. The mean vectors and covariance matrices are assumed

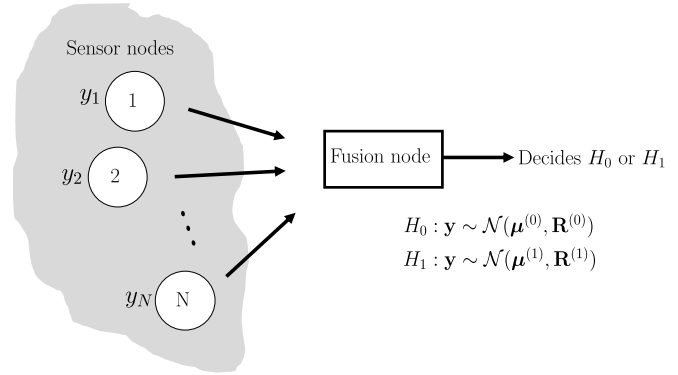


Fig. 1. A WSN consisting of N nodes that transmit their measurements to an FN, which decides in favor of a hypothesis.

to be known to the FN. The system model is illustrated in Fig. 1.

Optimum Decision Rule: Let c_{uv} be the cost incurred if hypothesis H_u is chosen when hypothesis H_v is true. Let ζ_0 and ζ_1 be the prior probabilities of H_0 and H_1 , respectively. The decision rule that minimizes the error probability is given by [29, Ch. III.A]

$$\log \left(\frac{f(\mathbf{y}|H_1)}{f(\mathbf{y}|H_0)} \right) \underset{H_0}{\overset{H_1}{\gtrless}} \beta, \quad (2)$$

where $f(\mathbf{y}|H_h)$ is the probability density function (PDF) of the measurements conditioned on the hypothesis H_h , and $\beta = \log(\frac{(c_{10} - c_{00})\zeta_0}{(c_{01} - c_{11})\zeta_1})$. For the signal model in (1), the decision rule in (2) becomes

$$\begin{aligned} &(\mathbf{y} - \boldsymbol{\mu}^{(0)})^T (\mathbf{R}^{(0)})^{-1} (\mathbf{y} - \boldsymbol{\mu}^{(0)}) + \log \left(\frac{\det(\mathbf{R}^{(0)})}{\det(\mathbf{R}^{(1)})} \right) \\ &- (\mathbf{y} - \boldsymbol{\mu}^{(1)})^T (\mathbf{R}^{(1)})^{-1} (\mathbf{y} - \boldsymbol{\mu}^{(1)}) \underset{H_0}{\overset{H_1}{\gtrless}} 2\beta. \end{aligned} \quad (3)$$

Expanding the products and rearranging terms yields

$$\begin{aligned} \mathbf{y}^T \mathbf{G} \mathbf{y} + \mathbf{s}^T \mathbf{y} - \log \left(\frac{\det(\mathbf{R}^{(1)})}{\det(\mathbf{R}^{(0)})} \right) - (\boldsymbol{\mu}^{(1)})^T (\mathbf{R}^{(1)})^{-1} \boldsymbol{\mu}^{(1)} \\ + (\boldsymbol{\mu}^{(0)})^T (\mathbf{R}^{(0)})^{-1} \boldsymbol{\mu}^{(0)} \underset{H_0}{\overset{H_1}{\gtrless}} 2\beta, \end{aligned} \quad (4)$$

where

$$\begin{aligned} \mathbf{G} &= (\mathbf{R}^{(0)})^{-1} - (\mathbf{R}^{(1)})^{-1}, \\ \mathbf{s} &= 2(\mathbf{R}^{(1)})^{-1} \boldsymbol{\mu}^{(1)} - 2(\mathbf{R}^{(0)})^{-1} \boldsymbol{\mu}^{(0)}. \end{aligned} \quad (5)$$

We see that cross-terms of the form $G_{ij}y_i y_j$ arise in (4) whenever \mathbf{G} is not a diagonal matrix.

The matrices $\mathbf{R}^{(0)}$ and $\mathbf{R}^{(1)}$ are assumed to be positive definite. If either of them is positive semi-definite, then the measurement of a node can be expressed as a linear combination of other nodes' measurements. In that case, the FN modifies the decision rule in (4) to only account for the measurements whose covariance sub-matrix is positive definite.

Timer Scheme: At the start of each measurement round, the FN broadcasts a beacon signal to all the nodes.

Upon receiving the beacon, node i sets a timer that is a monotone non-increasing function of a locally-computed real number called *metric*. All the nodes use the same metric-to-timer mapping. Once its timer expires, the node transmits a packet to the FN. The metric and the information contained in the packet are functions of the measurement y_i . We specify them in the next section. The timer scheme ensures that the nodes transmit in the decreasing order of their metrics, without any communication between the nodes. Its design can take into account several aspects such as the differences in the propagation and detection delays, and even physical layer capabilities of the nodes such as carrier sensing [17].

When the FN receives a packet from a node, it decides on a hypothesis or it waits for the next packet. Once the FN makes a decision, it broadcasts a control signal to all the sensor nodes to halt their timers for the rest of the round. The process starts afresh at the beginning of every round when the nodes make a new set of measurements. As in OTS, we assume that the quantization error is negligible and the FN receives with a negligible probability of decoding error and collision [6], [7], [9], [15], [16].

III. CORRELATION-AWARE OTS

We present CA-OTS for the general case in which both mean vectors and covariance matrices of the two hypotheses can be different. Thereafter, we present novel refinements for the following two specializations, which are important in their own right:

- *Shift-in-Covariance*: Here, $\mathbf{R}^{(0)} \neq \mathbf{R}^{(1)}$ but $\boldsymbol{\mu}^{(0)} = \boldsymbol{\mu}^{(1)}$. This happens, for example, when the nodes detect a Gaussian signal in the presence of spatially correlated Gaussian noise [29, Ch. III-B].
- *Shift-in-Mean*: In this, $\boldsymbol{\mu}^{(0)} \neq \boldsymbol{\mu}^{(1)}$ but $\mathbf{R}^{(0)} = \mathbf{R}^{(1)}$. This happens, for example, when the nodes detect a deterministic signal in the presence of spatially correlated Gaussian noise [29, Ch. III-B].

A. General Case

In CA-OTS, we specify that the metric of node i is $|y_i|$. Thus, node i sets its timer to be a monotone non-increasing function of $|y_i|$. Once the timer expires, it transmits a packet containing its measurement y_i and its index i . Thus, the FN receives measurements in the order $|y_{[1]}|, |y_{[2]}|, \dots, |y_{[N]}|$, where the subscript $[i]$ is the index of the node with the i^{th} largest metric. Expressing (4) in terms of the ordered statistics notation and rearranging the terms yields

$$d(\mathbf{y}) \triangleq \sum_{i=1}^N G_{[i][i]} y_{[i]}^2 + \sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N G_{[i][j]} y_{[i]} y_{[j]} + \sum_{i=1}^N s_{[i]} y_{[i]} \quad (6)$$

$$\underset{H_0}{\geq} \underset{H_1}{\beta_{\text{gen}}},$$

where $\beta_{\text{gen}} = 2\beta + (\boldsymbol{\mu}^{(1)})^T (\mathbf{R}^{(1)})^{-1} \boldsymbol{\mu}^{(1)} - (\boldsymbol{\mu}^{(0)})^T (\mathbf{R}^{(0)})^{-1} \boldsymbol{\mu}^{(0)} + \log(\det(\mathbf{R}^{(1)})/\det(\mathbf{R}^{(0)}))$.

When the FN has received measurements from the nodes $[1], [2], \dots, [k]$, we can write (6) in a compact matrix

form, which is as follows. Let $\mathbf{r}_k = [y_{[1]}, y_{[2]}, \dots, y_{[k]}]^T$ and $\mathbf{u}_{N-k} = [y_{[k+1]}, y_{[k+2]}, \dots, y_{[N]}]^T$ be the vector of measurements that are received and yet to be received, respectively. Let $\mathbf{P}_k = [G_{[i][j]}]$, for $1 \leq i, j \leq k$, $\mathbf{m}_k = [s_{[1]}, s_{[2]}, \dots, s_{[k]}]^T$, $\mathbf{T}_{N-k} = [G_{[i][j]}]$, for $k+1 \leq i, j \leq N$, and $\mathbf{v}_{N-k} = [(s_{[k+1]}/2) + \sum_{i=1}^k G_{[i][k+1]} y_{[i]}, (s_{[k+2]}/2) + \sum_{i=1}^k G_{[i][k+2]} y_{[i]}, \dots, (s_{[N]}/2) + \sum_{i=1}^k G_{[i][N]} y_{[i]}]^T$.

We see that \mathbf{P}_k is the sub-matrix of \mathbf{G} that corresponds to the k received measurements, \mathbf{m}_k consists of the entries of \mathbf{s} that correspond to the received measurements, \mathbf{T}_{N-k} is the sub-matrix of \mathbf{G} that corresponds to the $(N-k)$ nodes that have not yet transmitted, and the elements of \mathbf{v}_{N-k} depend on the nodes $[k+1], \dots, [N]$ that are yet to transmit and also the k received measurements. Thus, $d(\mathbf{y})$ in (6) can be recast as

$$d(\mathbf{y}) = \mathbf{r}_k^T \mathbf{P}_k \mathbf{r}_k + \mathbf{m}_k^T \mathbf{r}_k + \mathbf{u}_{N-k}^T \mathbf{T}_{N-k} \mathbf{u}_{N-k} + 2\mathbf{v}_{N-k}^T \mathbf{u}_{N-k}. \quad (7)$$

The FN can compute the terms $\mathbf{r}_k^T \mathbf{P}_k \mathbf{r}_k$ and $\mathbf{m}_k^T \mathbf{r}_k$ in (7) from the k measurements it has received. However, it does not know \mathbf{u}_{N-k} . It also does not know \mathbf{T}_{N-k} and \mathbf{v}_{N-k} since they depend on the order in which the remaining nodes will transmit. We explain this with an example.

Example: Let $N = 4$, $k = 2$, and $[1] = 2$ and $[2] = 4$. Thus, node 2 transmitted first followed by node 4. The FN can compute the term $\mathbf{r}_k^T \mathbf{P}_k \mathbf{r}_k + \mathbf{m}_k^T \mathbf{r}_k$. The FN also knows that nodes 1 and 3 have not transmitted. However, it does not know if $[3]$ is 1 or 3. It, thus, does not know \mathbf{T}_{N-k} and \mathbf{v}_{N-k} . As a result, it cannot compute the terms $\mathbf{u}_{N-k}^T \mathbf{T}_{N-k} \mathbf{u}_{N-k}$ and $\mathbf{v}_{N-k}^T \mathbf{u}_{N-k}$. However, notice that \mathbf{T}_{N-k} and \mathbf{v}_{N-k} when $[3] = 1$ and $[4] = 3$ are permuted versions of \mathbf{T}_{N-k} and \mathbf{v}_{N-k} when $[3] = 3$ and $[4] = 1$. We formalize this insight below.

Let $\ell_1, \ell_2, \dots, \ell_{N-k}$ denote the indices of the $(N-k)$ nodes that have not transmitted such that $\ell_1 < \ell_2 < \dots < \ell_{N-k}$. Let $\tilde{\mathbf{u}}_{N-k} = [y_{\ell_1}, y_{\ell_2}, \dots, y_{\ell_{N-k}}]^T$, $\tilde{\mathbf{T}}_{N-k} = [G_{\ell_i \ell_j}]$, for $1 \leq i, j \leq N-k$, and $\tilde{\mathbf{v}}_{N-k} = [(s_{\ell_1}/2) + \sum_{i=1}^k G_{[i]\ell_1} y_{[i]}, (s_{\ell_2}/2) + \sum_{i=1}^k G_{[i]\ell_2} y_{[i]}, \dots, (s_{\ell_{N-k}}/2) + \sum_{i=1}^k G_{[i]\ell_{N-k}} y_{[i]}]^T$. Here, $\ell_1, \ell_2, \dots, \ell_{N-k}$ can be interpreted as the *unordered indices*. In the example above, $\ell_1 = 1$ and $\ell_2 = 3$. Similarly, $\tilde{\mathbf{u}}_{N-k}$ can be interpreted as the vector of unordered yet-to-be-received measurements and $\tilde{\mathbf{T}}_{N-k}$ as the sub-matrix of \mathbf{G} that corresponds to these unordered measurements. Depending on the measurements, some permutation of them yields $[k+1], [k+2], \dots, [N]$. Let \mathbf{U}_π denote the permutation matrix of size $(N-k) \times (N-k)$ for this permutation. Then, $\mathbf{u}_{N-k} = \mathbf{U}_\pi \tilde{\mathbf{u}}_{N-k}$, $\mathbf{T}_{N-k} = \mathbf{U}_\pi \tilde{\mathbf{T}}_{N-k} \mathbf{U}_\pi^T$, and $\mathbf{v}_{N-k} = \mathbf{U}_\pi \tilde{\mathbf{v}}_{N-k}$. Using $\mathbf{U}_\pi \mathbf{U}_\pi^T = \mathbf{U}_\pi^T \mathbf{U}_\pi = \mathbf{I}_{N-k}$, (7) becomes

$$d(\mathbf{y}) = \mathbf{r}_k^T \mathbf{P}_k \mathbf{r}_k + \mathbf{m}_k^T \mathbf{r}_k + \tilde{\mathbf{u}}_{N-k}^T \tilde{\mathbf{T}}_{N-k} \tilde{\mathbf{u}}_{N-k} + 2\tilde{\mathbf{v}}_{N-k}^T \tilde{\mathbf{u}}_{N-k}. \quad (8)$$

Since the quantities $\tilde{\mathbf{T}}_{N-k}$ and $\tilde{\mathbf{v}}_{N-k}$ depend on the nodes that are yet to transmit and not on their ordering, the FN

can explicitly compute them. This is unlike \mathbf{T}_{N-k} and \mathbf{v}_{N-k} in (7). Furthermore, $|y_{\ell_i}| < |y_{[k]}|$, for $1 \leq i \leq N-k$. We can now bound the unknown terms $\tilde{\mathbf{u}}_{N-k}^T \tilde{\mathbf{T}}_{N-k} \tilde{\mathbf{u}}_{N-k}$ and $\tilde{\mathbf{v}}_{N-k}^T \tilde{\mathbf{u}}_{N-k}$ and, in effect, $d(\mathbf{y})$ in terms of the measurements received thus far. It leads to the following novel decision rules, which are derived in Appendix A.

Result 1: Let the FN have received $y_{[1]}, y_{[2]}, \dots, y_{[k]}$ from the nodes $[1], [2], \dots, [k]$ and

$$\mathbf{a}_k = \mathbf{r}_k^T \mathbf{P}_k \mathbf{r}_k + \mathbf{m}_k^T \mathbf{r}_k$$

$$\mathbf{b}_k = 2 |y_{[k]}| \sum_{j \in \mathcal{N} \setminus \{[1], [2], \dots, [k]\}} \left| \frac{s_j}{2} + \sum_{i=1}^k G_{[i]j} y_{[i]} \right|.$$

The following decision rules achieve the same optimal error probability as UTS:

1) If $\lambda_{\min}(\tilde{\mathbf{T}}_{N-k}) \geq 0$:

$$\text{Decide } H_1 \text{ if: } \mathbf{a}_k > \beta_{\text{gen}} + \mathbf{b}_k, \quad (9)$$

$$\text{Decide } H_0 \text{ if: } \mathbf{a}_k < \beta_{\text{gen}} - \lambda_{\max}(\tilde{\mathbf{T}}_{N-k}) (N-k) y_{[k]}^2 - \mathbf{b}_k, \quad (10)$$

Wait for the next transmission, otherwise.

2) If $\lambda_{\min}(\tilde{\mathbf{T}}_{N-k}) < 0 < \lambda_{\max}(\tilde{\mathbf{T}}_{N-k})$:

$$\text{Decide } H_1 \text{ if: } \mathbf{a}_k > \beta_{\text{gen}} - \lambda_{\min}(\tilde{\mathbf{T}}_{N-k}) (N-k) y_{[k]}^2 + \mathbf{b}_k, \quad (11)$$

$$\text{Decide } H_0 \text{ if: } \mathbf{a}_k < \beta_{\text{gen}} - \lambda_{\max}(\tilde{\mathbf{T}}_{N-k}) (N-k) y_{[k]}^2 - \mathbf{b}_k, \quad (12)$$

Wait for the next transmission, otherwise.

3) If $\lambda_{\max}(\tilde{\mathbf{T}}_{N-k}) \leq 0$:

$$\text{Decide } H_1 \text{ if: } \mathbf{a}_k > \beta_{\text{gen}} - \lambda_{\min}(\tilde{\mathbf{T}}_{N-k}) (N-k) y_{[k]}^2 + \mathbf{b}_k, \quad (13)$$

$$\text{Decide } H_0 \text{ if: } \mathbf{a}_k < \beta_{\text{gen}} - \mathbf{b}_k, \quad (14)$$

Wait for the next transmission, otherwise.

If $k = N$, the FN decides based on (6). ■

Remark: The FN can compute the lower and upper thresholds in the decision rules above since they are entirely in terms of the k measurements $y_{[1]}, y_{[2]}, \dots, y_{[k]}$ and the k node indices $[1], [2], \dots, [k]$ that it has received. For example, to compute the sum $\sum_{j \in \mathcal{N} \setminus \{[1], [2], \dots, [k]\}} |(s_j/2) + \sum_{i=1}^k G_{[i]j} y_{[i]}|$, the FN only needs to know who the remaining nodes are and not the order in which they will transmit.

B. Detecting a Shift in Covariance

In this case, $\mathbf{R}^{(0)} \neq \mathbf{R}^{(1)}$ but $\boldsymbol{\mu}^{(0)} = \boldsymbol{\mu}^{(1)} = \boldsymbol{\mu}$. We present two approaches below that bring out the important role that the metric plays in the design of the scheme.

We first write the decision rule in (4) as follows:

$$d_{\text{cov}}(\mathbf{z}) \triangleq \mathbf{z}^T \mathbf{G} \mathbf{z} \underset{H_0}{\underset{H_1}{\geq}} \beta_{\text{cov}}, \quad (15)$$

where $\beta_{\text{cov}} = 2\beta + \log(\det(\mathbf{R}^{(1)})/\det(\mathbf{R}^{(0)}))$ and $\mathbf{z} = \mathbf{y} - \boldsymbol{\mu}$. Expanding (15), we get

$$\sum_{i=1}^N G_{ii} z_i^2 + \sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N G_{ij} z_i z_j \underset{H_0}{\underset{H_1}{\geq}} \beta_{\text{cov}}. \quad (16)$$

We shall refer to $z_i = y_i - \mu_i$ as the mean-shifted measurement of node i .

We propose that the metric for node i is $|z_i|$. When its timer expires, it sends z_i and its identity i to the FN.¹ Now, $|z_{[1]}| > |z_{[2]}| > \dots > |z_{[N]}|$. Based on (15), we present two approaches to derive the decision rules, namely, covariance-shift-OTS (CovShift-OTS) and CovShift-Refined-OTS. They present different trade-offs between complexity and the number of transmissions.

1) *CovShift-OTS:* Here, we bound the quadratic term $\mathbf{z}^T \mathbf{G} \mathbf{z}$ to obtain the following decision rules. The derivation is given in Appendix B.

Result 2: Let the FN have received the mean-shifted measurements $z_{[1]}, z_{[2]}, \dots, z_{[k]}$. The following decision rules achieve the same error probability as UTS:

1) If $\lambda_{\min}(\mathbf{G}) \geq 0$:

$$\text{Decide } H_1 \text{ if: } \lambda_{\min}(\mathbf{G}) \sum_{i=1}^k z_{[i]}^2 > \beta_{\text{cov}}, \quad (17)$$

$$\text{Decide } H_0 \text{ if: } \lambda_{\max}(\mathbf{G}) \sum_{i=1}^k z_{[i]}^2 < \beta_{\text{cov}} - \lambda_{\max}(\mathbf{G}) (N-k) z_{[k]}^2, \quad (18)$$

Wait for the next transmission, otherwise.

2) If $\lambda_{\min}(\mathbf{G}) < 0 < \lambda_{\max}(\mathbf{G})$:

$$\text{Decide } H_1 \text{ if: } \lambda_{\min}(\mathbf{G}) \sum_{i=1}^k z_{[i]}^2 > \beta_{\text{cov}} - \lambda_{\min}(\mathbf{G}) (N-k) z_{[k]}^2, \quad (19)$$

$$\text{Decide } H_0 \text{ if: } \lambda_{\max}(\mathbf{G}) \sum_{i=1}^k z_{[i]}^2 < \beta_{\text{cov}} - \lambda_{\max}(\mathbf{G}) (N-k) z_{[k]}^2, \quad (20)$$

Wait for the next transmission, otherwise.

3) If $\lambda_{\max}(\mathbf{G}) \leq 0$:

$$\text{Decide } H_1 \text{ if: } \lambda_{\min}(\mathbf{G}) \sum_{i=1}^k z_{[i]}^2 > \beta_{\text{cov}} - \lambda_{\min}(\mathbf{G}) (N-k) z_{[k]}^2, \quad (21)$$

$$\text{Decide } H_0 \text{ if: } \lambda_{\max}(\mathbf{G}) \sum_{i=1}^k z_{[i]}^2 < \beta_{\text{cov}}, \quad (22)$$

Wait for the next transmission, otherwise.

If $k = N$, the FN decides based on (16). ■

¹For this, a node i is assumed to know its mean μ_i , which is easy to ensure in practice.

Analytical Bounds: To understand the performance of CovShift-OTS, we derive an upper bound on the average number of transmissions \bar{A} . While this is intractable for arbitrary $\mathbf{R}^{(0)}$ and $\mathbf{R}^{(1)}$, we derive bounds for the product-correlation model [22]. In it, under H_1 , the correlation-coefficient between the measurements y_i and y_j at nodes i and j is $\theta_i\theta_j$, where $\theta_1, \theta_2, \dots, \theta_N \in (-1, 1)$. Thus, $R_{ij}^{(1)} = \sigma_i\sigma_j\theta_i\theta_j$, for $i \neq j$, and σ_i^2 , otherwise, where σ_i^2 is the variance of the measurement at node i under H_1 . Note that the variances $\sigma_1^2, \sigma_2^2, \dots, \sigma_N^2$ can be different. Furthermore, $\mathbf{R}^{(0)} = \gamma_0 \mathbf{I}_N$. Thus, only spatially white noise is measured under H_0 [16]. From the definition of \mathbf{G} in (5), we get

$$\begin{aligned} \lambda_{\min}(\mathbf{G}) &= \frac{1}{\gamma_0} - \frac{1}{\lambda_{\min}(\mathbf{R}^{(1)})}, \\ \lambda_{\max}(\mathbf{G}) &= \frac{1}{\gamma_0} - \frac{1}{\lambda_{\max}(\mathbf{R}^{(1)})}. \end{aligned} \quad (23)$$

Let

$$\alpha_1 = \frac{\beta_{\text{cov}}}{\lambda_{\min}(\mathbf{G})} \quad \text{and} \quad \alpha_0 = \frac{\beta_{\text{cov}}}{\lambda_{\max}(\mathbf{G})}. \quad (24)$$

Result 3: \bar{A} is upper bounded by

$$\bar{A} \leq N - (N - 1) \bar{\nu}, \quad (25)$$

where $\bar{\nu}$ depends on the signs of β_{cov} , $\lambda_{\min}(\mathbf{G})$, and $\lambda_{\max}(\mathbf{G})$. It is given as follows:

1) If $\lambda_{\min}(\mathbf{G}) \geq 0$:

$$\bar{\nu} = \begin{cases} 1, & \text{if } \beta_{\text{cov}} < 0, \\ \begin{aligned} & \left[1 - \delta_0(\alpha_1) + \delta_0\left(\frac{\alpha_0}{N}\right) \right] \zeta_0 \\ & + \left[1 - \delta_1(\alpha_1) + \delta_1\left(\frac{\alpha_0}{N}\right) \right] \zeta_1, \end{aligned} & \text{otherwise.} \end{cases} \quad (26)$$

2) If $\lambda_{\min}(\mathbf{G}) < 0 < \lambda_{\max}(\mathbf{G})$:

$$\bar{\nu} = \begin{cases} \delta_0\left(\frac{\alpha_1}{N}\right) \zeta_0 + \delta_1\left(\frac{\alpha_1}{N}\right) \zeta_1, & \text{if } \beta_{\text{cov}} < 0, \\ \delta_0\left(\frac{\alpha_0}{N}\right) \zeta_0 + \delta_1\left(\frac{\alpha_0}{N}\right) \zeta_1, & \text{otherwise.} \end{cases} \quad (27)$$

3) If $\lambda_{\max}(\mathbf{G}) \leq 0$:

$$\bar{\nu} = \begin{cases} \begin{aligned} & \left[1 - \delta_0(\alpha_0) + \delta_0\left(\frac{\alpha_1}{N}\right) \right] \zeta_0 \\ & + \left[1 - \delta_1(\alpha_0) + \delta_1\left(\frac{\alpha_1}{N}\right) \right] \zeta_1, \end{aligned} & \text{if } \beta_{\text{cov}} < 0, \\ 1, & \text{otherwise,} \end{cases} \quad (28)$$

where $\delta_0(x) = \left[1 - 2Q\left(\sqrt{\frac{x}{\gamma_0}}\right) \right]^N$,

$$\begin{aligned} \delta_1(x) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \prod_{i=1}^N \left[Q\left(\frac{1}{\sqrt{1-\theta_i^2}} \left[-\sqrt{\frac{x}{\sigma_i^2}} - \theta_i\psi \right] \right) \right. \\ &\quad \left. - Q\left(\frac{1}{\sqrt{1-\theta_i^2}} \left[\sqrt{\frac{x}{\sigma_i^2}} - \theta_i\psi \right] \right) \right] e^{-\frac{\psi^2}{2}} d\psi, \end{aligned} \quad (29)$$

and $Q(\cdot)$ denotes the Gaussian Q -function [30, Ch. 26.2].

Proof: The proof is given in Appendix C. ■

Remark: The different cases arise in Result 3 because β_{cov} , $\lambda_{\min}(\mathbf{G})$, and $\lambda_{\max}(\mathbf{G})$ can be positive or negative. The integral in (29) can be easily evaluated using Gauss-Hermite quadrature [30, Ch. 25.4] as follows:

$$\begin{aligned} \delta_1(x) &\approx \sum_{l=1}^g \frac{\omega_l}{\sqrt{\pi}} \prod_{i=1}^N \left[Q\left(\frac{1}{\sqrt{1-\theta_i^2}} \left[-\sqrt{\frac{x}{\sigma_i^2}} - \sqrt{2}\theta_i a_l \right] \right) \right. \\ &\quad \left. - Q\left(\frac{1}{\sqrt{1-\theta_i^2}} \left[\sqrt{\frac{x}{\sigma_i^2}} - \sqrt{2}\theta_i a_l \right] \right) \right], \end{aligned}$$

where ω_l and a_l , for $1 \leq l \leq g$, are the Gauss-Hermite weights and abscissas, respectively, and g is the number of terms.²

To gain insights, consider the case where γ_0 is fixed and $\sigma_1^2 = \sigma_2^2 = \dots = \sigma_N^2 = \gamma_1$. Define γ_1/γ_0 as the SNR. When $\gamma_1 \rightarrow \infty$ and γ_0 is fixed, $\text{SNR} \rightarrow \infty$. It follows from (23) that $\lambda_{\min}(\mathbf{G}) \rightarrow 1/\gamma_0$, $\lambda_{\max}(\mathbf{G}) \rightarrow 1/\gamma_0$, and $\beta_{\text{cov}} \rightarrow \infty$. Thus, $\delta_0(\alpha_1) \rightarrow 1$ and $\delta_0(\alpha_0/N) \rightarrow 1$. Furthermore, $\delta_1(\alpha_1) \rightarrow 0$ and $\delta_1(\alpha_1/N) \rightarrow 0$. Thus, $\bar{\nu} \rightarrow 1$ and $\bar{A} \rightarrow 1$. Hence, on an average, the FN can make a decision with a single transmission without any increase in the error probability.

We see from Result 2 that the FN does not need to know the identities of the nodes that have transmitted to implement the decision rules for CovShift-OTS, which is unlike CA-OTS.

2) *CovShift-Refined-OTS:* As in CovShift-OTS, the metric of node i is $|z_i|$ and it transmits z_i upon timer expiry. Thus, as before, $|z_{[1]}| > |z_{[2]}| > \dots > |z_{[N]}|$.

In terms of the order statistics notation, the decision statistic $d_{\text{cov}}(\mathbf{z})$ can be expressed as

$$d_{\text{cov}}(\mathbf{z}) = \sum_{i=1}^N G_{[i][i]} z_{[i]}^2 + \sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N G_{[i][j]} z_{[i]} z_{[j]}. \quad (30)$$

When the FN has received the mean-shifted measurements $z_{[1]}, z_{[2]}, \dots, z_{[k]}$ from the nodes $[1], [2], \dots, [k]$, we can express (30) in the following form that is similar to (7):

$$\begin{aligned} d_{\text{cov}}(\mathbf{z}) &= (\mathbf{r}'_k)^T \mathbf{P}_k \mathbf{r}'_k + (\mathbf{u}'_{N-k})^T \mathbf{T}_{N-k} \mathbf{u}'_{N-k} \\ &\quad + 2(\mathbf{v}'_{N-k})^T \mathbf{u}'_{N-k}, \end{aligned} \quad (31)$$

where \mathbf{P}_k and \mathbf{T}_{N-k} have been defined earlier, $\mathbf{r}'_k = [z_{[1]}, z_{[2]}, \dots, z_{[k]}]^T$, $\mathbf{u}'_{N-k} = [z_{[k+1]}, z_{[k+2]}, \dots, z_{[N]}]^T$, and $\mathbf{v}'_{N-k} = \left[\sum_{i=1}^k G_{[i][k+1]} z_{[i]}, \sum_{i=1}^k G_{[i][k+2]} z_{[i]}, \dots, \sum_{i=1}^k G_{[i][N]} z_{[i]} \right]^T$. As in Section III-A, the FN constructs $\tilde{\mathbf{T}}_{N-k}$, which is the sub-matrix of \mathbf{G} that corresponds to the unordered measurements from the nodes in the set $\mathbb{N} \setminus \{[1], [2], \dots, [k]\}$. The decision rules for CovShift-Refined-OTS are as follows.

Result 4: Let the FN have received the mean-shifted measurements $z_{[1]}, z_{[2]}, \dots, z_{[k]}$ from the nodes $[1], [2], \dots, [k]$ and

$$\begin{aligned} \mathbf{a}'_k &= (\mathbf{r}'_k)^T \mathbf{P}_k \mathbf{r}'_k, \\ \mathbf{b}'_k &= 2 |z_{[k]}| \sum_{j \in \mathbb{N} \setminus \{[1], [2], \dots, [k]\}} \left| \sum_{i=1}^k G_{[i]j} z_{[i]} \right|. \end{aligned}$$

²We have found that $g = 8$ is sufficient to ensure numerical accuracy for all system parameter values of interest.

The following decision rules achieve the same error probability as UTS:

1) If $\lambda_{\min}(\tilde{\mathbf{T}}_{N-k}) \geq 0$:

$$\text{Decide } H_1 \text{ if: } \mathbf{a}'_k > \beta_{\text{cov}} + \mathbf{b}'_k, \quad (32)$$

$$\text{Decide } H_0 \text{ if: } \mathbf{a}'_k < \beta_{\text{cov}} - \lambda_{\max}(\tilde{\mathbf{T}}_{N-k}) (N-k) z_{[k]}^2 - \mathbf{b}'_k, \quad (33)$$

Wait for the next transmission, otherwise.

2) If $\lambda_{\min}(\tilde{\mathbf{T}}_{N-k}) < 0 < \lambda_{\max}(\tilde{\mathbf{T}}_{N-k})$:

$$\text{Decide } H_1 \text{ if: } \mathbf{a}'_k > \beta_{\text{cov}} - \lambda_{\min}(\tilde{\mathbf{T}}_{N-k}) (N-k) z_{[k]}^2 + \mathbf{b}'_k, \quad (34)$$

$$\text{Decide } H_0 \text{ if: } \mathbf{a}'_k < \beta_{\text{cov}} - \lambda_{\max}(\tilde{\mathbf{T}}_{N-k}) (N-k) z_{[k]}^2 - \mathbf{b}'_k, \quad (35)$$

Wait for the next transmission, otherwise.

3) If $\lambda_{\max}(\tilde{\mathbf{T}}_{N-k}) \leq 0$:

$$\text{Decide } H_1 \text{ if: } \mathbf{a}'_k > \beta_{\text{cov}} - \lambda_{\min}(\tilde{\mathbf{T}}_{N-k}) (N-k) z_{[k]}^2 + \mathbf{b}'_k, \quad (36)$$

$$\text{Decide } H_0 \text{ if: } \mathbf{a}'_k < \beta_{\text{cov}} - \mathbf{b}'_k, \quad (37)$$

Wait for the next transmission, otherwise.

If $k = N$, the FN decides based on (16).

Proof: The proof is similar to that in Appendix A, and is skipped to conserve space. ■

The decision rules for CovShift-OTS use $\lambda_{\min}(\mathbf{G})$ and $\lambda_{\max}(\mathbf{G})$, which only needs to be computed once. However, this approach bounds the term $(\mathbf{r}'_k)^T \mathbf{P}_k \mathbf{r}'_k$ instead of computing it exactly, which CovShift-Refined-OTS does. The looser bounds increase the odds that the FN requires more transmissions to decide. CovShift-OTS is a refinement over CA-OTS as it uses a more refined metric. However, it is more involved than CovShift-OTS as it needs to compute the smallest and largest eigenvalues of $\tilde{\mathbf{T}}_{N-k}$, which changes every time the FN receives a measurement. This discussion also implies that the upper bound in Result 3 also applies to CovShift-Refined-OTS.

C. Detecting a Shift in Mean

In this case, $\boldsymbol{\mu}^{(0)} \neq \boldsymbol{\mu}^{(1)}$ but $\mathbf{R}^{(1)} = \mathbf{R}^{(0)} = \mathbf{R}$. Thus, the decision rule in (4) simplifies to

$$\mathbf{s}^T \mathbf{y} \underset{H_0}{\overset{H_1}{\geq}} \beta_{\text{mean}}, \quad (38)$$

where $\mathbf{s} = 2\mathbf{R}^{-1}(\boldsymbol{\mu}^{(1)} - \boldsymbol{\mu}^{(0)})$ and $\beta_{\text{mean}} = 2\beta + (\boldsymbol{\mu}^{(1)})^T \mathbf{R}^{-1} \boldsymbol{\mu}^{(1)} - (\boldsymbol{\mu}^{(0)})^T \mathbf{R}^{-1} \boldsymbol{\mu}^{(0)}$. Thus, (38) becomes

$$\sum_{i=1}^N w_i \underset{H_0}{\overset{H_1}{\geq}} \beta_{\text{mean}}, \quad (39)$$

where, for $1 \leq i \leq N$,

$$w_i = s_i y_i \text{ and } s_i = 2 \sum_{j=1}^N R_{ij}^{-1} (\mu_j^{(1)} - \mu_j^{(0)}), \quad (40)$$

and R_{ij}^{-1} denotes the element in the i^{th} row and j^{th} column of \mathbf{R}^{-1} .

In this case, the measurement y_i of node i is weighted by the node-specific scaling factor s_i that depends on the covariance matrix \mathbf{R} and the mean vectors $\boldsymbol{\mu}^{(0)}$ and $\boldsymbol{\mu}^{(1)}$. We propose the MeanShift-OTS scheme in which the metric of node i is $|w_i|$ and it transmits a packet containing w_i when its timer expires.³ Now, $|w_{[1]}| > |w_{[2]}| > \dots > |w_{[N]}|$. The decision rules for MeanShift-OTS are as follows.

Result 5: Let the FN have received $w_{[1]}, w_{[2]}, \dots, w_{[k]}$ from nodes $[1], [2], \dots, [k]$. The decision rules, which ensure the same error probability as UTS, are as follows:

$$\text{Decide } H_1 \text{ if: } \sum_{i=1}^k w_{[i]} > \beta_{\text{mean}} + (N-k) |w_{[k]}|, \quad (41)$$

$$\text{Decide } H_0 \text{ if: } \sum_{i=1}^k w_{[i]} < \beta_{\text{mean}} - (N-k) |w_{[k]}|, \quad (42)$$

Wait for the next transmission, otherwise.

Proof: The proof is similar to that of Theorem 1 in [6] and is skipped. ■

When $k = N$, the decision rule is the same as UTS, which is given in (39).

Analytical Bounds: To gain analytical insights, we consider the product-correlation model from before. In it, $R_{ij} = \sigma_i \sigma_j \theta_i \theta_j$, for $i \neq j$, $1 \leq i, j \leq N$, and $R_{ii} = \sigma_i^2$, for $1 \leq i \leq N$. And, $\boldsymbol{\mu}^{(0)} = \mathbf{0}$ [13], [15]. We assume the uniform cost model in which $c_{01} = c_{10} = 1$ and $c_{00} = c_{11} = 0$, and $\zeta_0 = \zeta_1 = 0.5$ [6], [16]. Thus, $\beta_{\text{mean}} = (\boldsymbol{\mu}^{(1)})^T \mathbf{R}^{-1} \boldsymbol{\mu}^{(1)}$. We now derive an upper bound on the average number of transmissions \bar{A} .

Result 6: \bar{A} is upper bounded as

$$\bar{A} \leq N - \left(\left\lfloor \frac{N}{2} \right\rfloor - 1 \right) [(\eta_0 + \kappa_0) \zeta_0 + (\eta_1 + \kappa_1) \zeta_1], \quad (43)$$

where the constants η_h and κ_h , for $h \in \{0, 1\}$, are given by

$$\eta_h = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \prod_{i=1}^N \left[Q \left(\frac{1}{\sqrt{1-\theta_i^2}} \left[\frac{\beta_{\text{mean}} - s_i \mu_i^{(h)}}{|s_i| \sigma_i} - \text{sgn}(s_i) \theta_i \psi \right] \right) \right] e^{-\frac{\psi^2}{2}} d\psi, \quad (44)$$

$$\kappa_h = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \prod_{i=1}^N \left[1 - Q \left(\frac{1}{\sqrt{1-\theta_i^2}} \left[\frac{\beta_{\text{mean}} - s_i \mu_i^{(h)}}{|s_i| \sigma_i} - \text{sgn}(s_i) \theta_i \psi \right] \right) \right] e^{-\frac{\psi^2}{2}} d\psi. \quad (45)$$

Proof: The proof is given in Appendix D. ■

As before, using Gauss-Hermite quadrature, the integrals in (44) and (45) can be easily computed as follows for $h \in \{0, 1\}$: $\eta_h \approx \sum_{i=1}^g \frac{\omega_i}{\sqrt{\pi}} \prod_{i=1}^N \left[Q \left(\frac{1}{\sqrt{1-\theta_i^2}} \left[\frac{\beta_{\text{mean}} - s_i \mu_i^{(h)}}{|s_i| \sigma_i} - \text{sgn}(s_i) \sqrt{2} \theta_i a_i \right] \right) \right]$

³This requires node i to know s_i . This can be communicated to the node by the FN, and it needs to be done only once.

$$\text{and } \kappa_h \approx \sum_{l=1}^g \frac{\omega_l}{\sqrt{\pi}} \prod_{i=1}^N \left[1 - Q \left(\frac{1}{\sqrt{1-\theta_i^2}} \left[\frac{\beta_{\text{mean}} - s_i \mu_i^{(h)}}{|s_i| \sigma_i} - \text{sgn}(s_i) \sqrt{2} \theta_i a_i \right] \right) \right].$$

To understand Result 5, consider $\mu_1^{(1)} = \mu_2^{(1)} = \dots = \mu_N^{(1)} = \mu$ and the uniform correlation model, in which $R_{ij} = \gamma\rho$, for $i \neq j$, $1 \leq i, j \leq N$, and $R_{ii} = \gamma$, for $1 \leq i \leq N$. It follows that $\theta_i = \sqrt{\rho}$ and $s_i = 2\mu/(\gamma[1 + \rho(N-1)])$, for $1 \leq i \leq N$. Hence, under H_1 , $(\beta_{\text{mean}} - s_i \mu_i^{(1)})/(|s_i| \sigma_i) = (N-2)\mu/(2\sqrt{\gamma})$, and under H_0 , $(\beta_{\text{mean}} - s_i \mu_i^{(0)})/(|s_i| \sigma_i) = N\mu/(2\sqrt{\gamma})$. Consequently, if $\mu \rightarrow \infty$ and γ is fixed, we have $\eta_0 \rightarrow 0$, $\eta_1 \rightarrow 0$, $\kappa_0 \rightarrow 1$, and $\kappa_1 \rightarrow 1$. And we get, $\bar{A} \leq \lceil N/2 \rceil + 1$.

IV. NUMERICAL RESULTS

We now benchmark the average number of transmissions \bar{A} since the error probability is the same as that of UTS. A reduction in \bar{A} also translates into an improvement in energy-efficiency as follows. The energy-efficiency of a scheme is the ratio of the average energy consumed by UTS to the average energy consumed by the scheme to make a decision. The larger this value is compared to one, the more energy-efficient the scheme is. When the nodes transmit with the same power, it follows that the energy-efficiency is equal to N/\bar{A} . When the nodes use different powers, e.g., when they are at different distances from the FN and employ power control, a reduction in \bar{A} implies a higher energy-efficiency, but the connection between the two is not as mathematically direct.

We illustrate results for the uniform cost model and $\zeta_0 = \zeta_1 = 0.5$. The Monte Carlo simulations are averaged over 10^6 measurement rounds. We show results for two correlation models. First, to demonstrate the generality of our model, we show results for the product-correlation model. In it, we set $\theta_i = \theta_{\text{max}} - (i-1)(\theta_{\text{max}} - \theta_{\text{min}})/(N-1)$, for $1 \leq i \leq N$, $\theta_{\text{min}} = 0.25$, and $\theta_{\text{max}} = 0.75$. Thus, the correlation between the measurements of different pairs of nodes is different. Second, to evaluate the impact of correlation, we show results for the uniform correlation model, in which the correlation-coefficient between any two nodes is $\rho \in (0, 1)$. The advantage of this model is that ρ is a single parameter that can be varied. Note that for these correlation models, there is only one cluster with a single cluster head as per the protocols proposed in [15] and [16]. Since all nodes within a cluster have to transmit to the cluster head, the number of transmissions for these protocols is N , which is the same as that of UTS.

A. General Case: Shift in Both Mean and Covariance

We show results for an example that involves both product-correlation and uniform correlation models. Specifically, $\mathbf{R}^{(1)}$ follows the product-correlation model with $R_{ij}^{(1)} = \theta_i \theta_j$, for $i \neq j$, $1 \leq i, j \leq N$, and $R_{ii}^{(1)} = 1$, for $1 \leq i \leq N$. And, $\mathbf{R}^{(0)}$ follows the uniform correlation model with $R_{ij}^{(0)} = 0.5$, for $i \neq j$, $1 \leq i, j \leq N$, and $R_{ii}^{(0)} = 1$, for $1 \leq i \leq N$. We set $\mu^{(0)} = \mathbf{0}$ and $\mu^{(1)} = \mu [1, 1, \dots, 1]^T$, and vary μ . Fig. 2 plots \bar{A} as a function of the number of sensors N for UTS and CA-OTS for different values of μ . For UTS, $\bar{A} = N$. As N

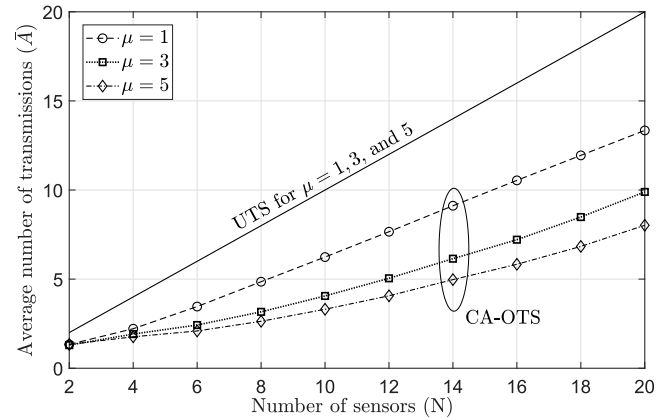


Fig. 2. Shift in both mean and covariance: \bar{A} as a function of N for different values of μ ($\mu^{(0)} = \mathbf{0}$, $\mu^{(1)} = \mu [1, 1, \dots, 1]^T$, $R_{ij}^{(0)} = 0.5$, for $i \neq j$, $1 \leq i, j \leq N$, $R_{ii}^{(0)} = 1$, for $1 \leq i \leq N$, $R_{ij}^{(1)} = \theta_i \theta_j$, for $i \neq j$, $1 \leq i, j \leq N$, and $R_{ii}^{(1)} = 1$, for $1 \leq i \leq N$).

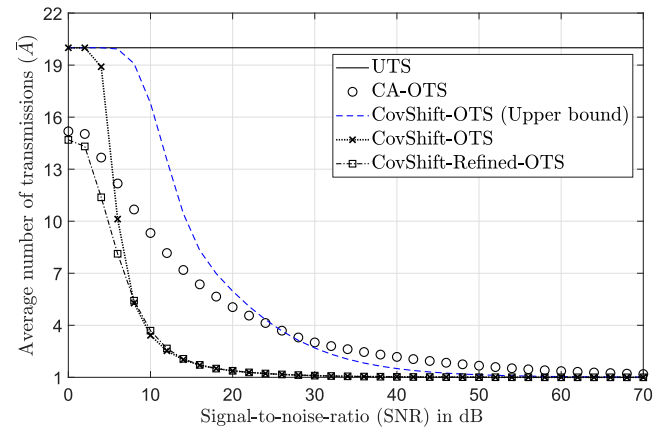


Fig. 3. Shift-in-covariance: \bar{A} as a function of the SNR γ_1/γ_0 for the product-correlation model ($N = 20$, $\mu^{(0)} = \mu^{(1)} = [1, 1, \dots, 1]^T$, $\mathbf{R}^{(0)} = \gamma_0 \mathbf{I}_N$, $R_{ij}^{(1)} = \gamma_1 \theta_i \theta_j$, for $i \neq j$, $1 \leq i, j \leq N$, and $R_{ii}^{(1)} = \gamma_1$, for $1 \leq i \leq N$).

increases, \bar{A} of CA-OTS increases. However, for all values of μ , it is much smaller than N . For example, when $N = 20$, CA-OTS requires 31%, 51%, and 59% fewer transmissions than UTS for $\mu = 1, 3$, and 5 , respectively. \bar{A} decreases as μ increases because the distributions become more separated.

B. Shift-in-Covariance

Product-Correlation Model: Fig. 3 plots \bar{A} of CA-OTS, CovShift-OTS, CovShift-Refined-OTS, and UTS for $N = 20$ nodes. For hypothesis H_1 , we set $R_{ij}^{(1)} = \gamma_1 \theta_i \theta_j$, for $i \neq j$, $1 \leq i, j \leq N$, and $R_{ii}^{(1)} = \gamma_1$, for $1 \leq i \leq N$. For hypothesis H_0 , we set $\mathbf{R}^{(0)} = \gamma_0 \mathbf{I}_N$. And, $\mu^{(0)} = \mu^{(1)} = [1, 1, \dots, 1]^T$. As the SNR γ_1/γ_0 increases, CA-OTS, CovShift-OTS, and CovShift-Refined-OTS all require fewer transmissions, on average, to make a decision compared to UTS. CovShift-Refined-OTS requires fewer transmissions than CA-OTS at all SNRs. For SNRs less than 6 dB, CA-OTS requires fewer transmissions than CovShift-OTS. This happens because the effect of cross-terms is more pronounced at smaller SNRs. CA-OTS computes the cross-terms that involve only the received measurements instead of bounding them,

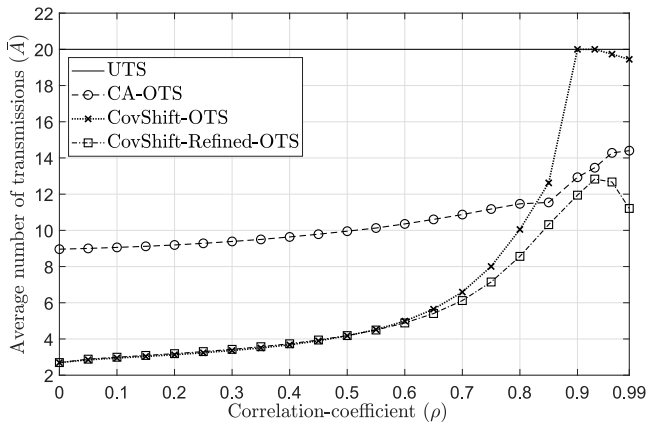


Fig. 4. Effect of correlation on the shift-in-covariance problem: \bar{A} as a function of ρ for the uniform correlation model ($N = 20$, $\boldsymbol{\mu}^{(0)} = \boldsymbol{\mu}^{(1)} = [1, 1, \dots, 1]^T$, $\mathbf{R}^{(0)} = \gamma_0 \mathbf{I}_N$, $R_{ij}^{(1)} = \gamma_1 \rho$, for $i \neq j$, $1 \leq i, j \leq N$, $R_{ii}^{(1)} = \gamma_1$, for $1 \leq i \leq N$, and $\gamma_1/\gamma_0 = 10$ dB).

as is done by CovShift-OTS. When the SNR increases, even CovShift-OTS requires fewer transmissions than CA-OTS because it uses a more refined metric. Also plotted is the upper bound on \bar{A} for CovShift-OTS from Result 3. It becomes tighter as the SNR increases and eventually becomes one.

Uniform Correlation Model: Fig. 4 plots \bar{A} of all the schemes as a function of the correlation-coefficient ρ for $\boldsymbol{\mu}^{(0)} = \boldsymbol{\mu}^{(1)} = [1, 1, \dots, 1]^T$, $\mathbf{R}^{(0)} = \gamma_0 \mathbf{I}_N$, $R_{ij}^{(1)} = \gamma_1 \rho$, for $i \neq j$, $1 \leq i, j \leq N$, and $R_{ii}^{(1)} = \gamma_1$, for $1 \leq i \leq N$. This is done for $N = 20$ and $\gamma_1/\gamma_0 = 10$ dB. As before, the ordered transmissions schemes require fewer transmissions than UTS, and CovShift-Refined-OTS requires even fewer transmissions than CA-OTS for all values of ρ . Even CovShift-OTS, despite its simplicity, requires fewer transmissions than CA-OTS for $\rho \leq 0.8$. For example, when $\rho = 0.4$, CA-OTS, CovShift-OTS, and CovShift-Refined-OTS require 50%, 80%, and 80%, respectively, fewer transmissions than UTS. However, when $\rho > 0.8$, the effect of cross-terms increases and CovShift-OTS requires more transmissions than CA-OTS.

As ρ increases from 0 to 0.9, \bar{A} for CovShift-OTS and CovShift-Refined-OTS increases. It stays the same until $\rho = 0.93$. However, for $\rho > 0.93$, it decreases. This can be understood using the notion of generalized variance. It is a measure of dispersion for a multivariate distribution and is equal to the determinant of the covariance matrix [31, Ch. 7.5]. This makes it equivalent to entropy, which is an affine function of the above determinant for Gaussian statistics. From [20], $\det(\mathbf{R}^{(1)}) = \gamma_1^N (1 - \rho)^{N-1} (1 + (N-1)\rho)$. And, $\det(\mathbf{R}^{(0)}) = \gamma_0^N$. At $\rho = 0$, $\det(\mathbf{R}^{(1)}) > \det(\mathbf{R}^{(0)})$. Intuitively, the more the difference between the two dispersions, the easier it is to distinguish between the two hypotheses. As ρ increases, $\det(\mathbf{R}^{(1)})$ decreases and it becomes more difficult to distinguish H_1 from H_0 . Thus, \bar{A} increases. When $\rho = 0.9$, \bar{A} of CovShift-OTS is the same as UTS. Here, it turns out that $\lambda_{\min}(\mathbf{G}) = 0$ and $\beta_{\text{cov}} > 0$. Consequently, the condition for deciding H_0 early is never satisfied in (17). Each additional measurement is less informative when ρ is large. This implies that the FN is also often unable to decide early about H_1 . When $\rho > 0.93$, $\det(\mathbf{R}^{(1)})$ becomes significantly smaller than

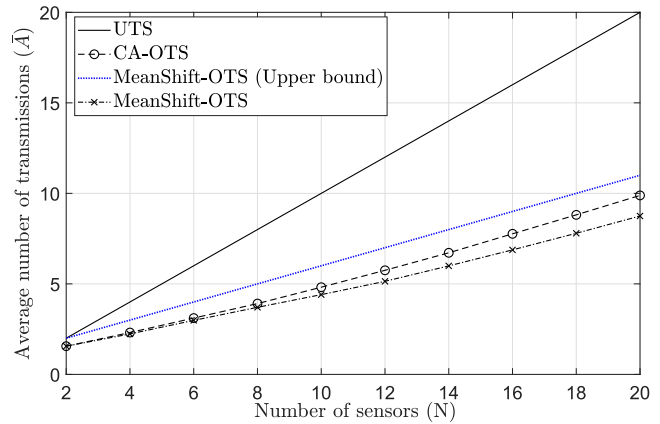


Fig. 5. Shift-in-mean: \bar{A} as a function of N for the product-correlation model ($\boldsymbol{\mu}^{(0)} = \mathbf{0}$, $\boldsymbol{\mu}^{(1)} = [3, 3, \dots, 3]^T$, $R_{ij}^{(0)} = R_{ij}^{(1)} = \theta_i \theta_j$, for $i \neq j$, $1 \leq i, j \leq N$, and $R_{ii}^{(0)} = R_{ii}^{(1)} = 1$, for $1 \leq i \leq N$).

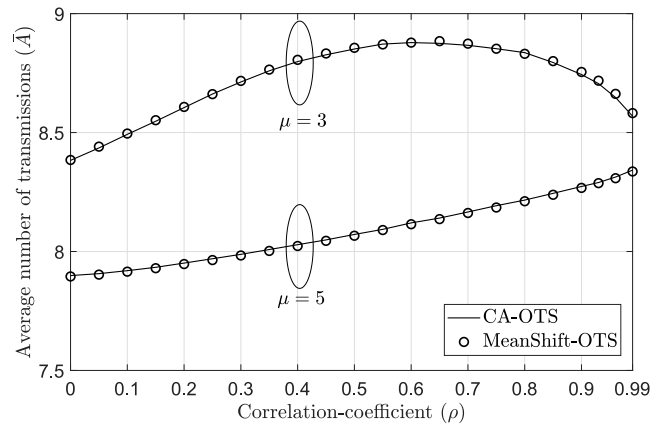


Fig. 6. Effect of correlation on the shift-in-mean problem: Zoomed-in-view of \bar{A} as a function of ρ for different values of μ for MeanShift-OTS and CA-OTS for the uniform correlation model ($N = 20$, $\boldsymbol{\mu}^{(0)} = \mathbf{0}$, $\boldsymbol{\mu}^{(1)} = \mu [1, 1, \dots, 1]^T$, $R_{ij}^{(0)} = R_{ij}^{(1)} = \rho$, for $i \neq j$, $1 \leq i, j \leq N$, and $R_{ii}^{(0)} = R_{ii}^{(1)} = 1$, for $1 \leq i \leq N$).

$\det(\mathbf{R}^{(0)})$ and it is easier to distinguish between H_1 and H_0 . Thus, \bar{A} decreases.

C. Shift-in-Mean

Product-Correlation Model: Fig. 5 plots \bar{A} of CA-OTS, MeanShift-OTS, and UTS as a function of N . This is done for $\boldsymbol{\mu}^{(0)} = \mathbf{0}$, $\boldsymbol{\mu}^{(1)} = [3, 3, \dots, 3]^T$, and the covariance matrix for both hypotheses follows the product-correlation model with $R_{ij}^{(0)} = R_{ij}^{(1)} = \theta_i \theta_j$, for $i \neq j$, $1 \leq i, j \leq N$, and $R_{ii}^{(0)} = R_{ii}^{(1)} = 1$, for $1 \leq i \leq N$. As N increases, the average number of transmissions of CA-OTS and MeanShift-OTS increase. Both CA-OTS and MeanShift-OTS have a much smaller \bar{A} than UTS. Also plotted is the upper bound on \bar{A} that is derived in Result 6. It tracks the simulation curve well.

Uniform Correlation Model: To understand the effect of correlation, Fig. 6 plots \bar{A} as a function of ρ for MeanShift-OTS and CA-OTS. This is done for $N = 20$, $\boldsymbol{\mu}^{(0)} = \mathbf{0}$, $\boldsymbol{\mu}^{(1)} = \mu [1, 1, \dots, 1]^T$, and $R_{ij}^{(0)} = R_{ij}^{(1)} = \rho$, for $i \neq j$, $1 \leq i, j \leq N$, and $R_{ii}^{(0)} = R_{ii}^{(1)} = 1$, for $1 \leq i \leq N$. We show results for $\mu = 3$ and 5. The results for UTS are not shown given its

substantially worse performance. Both the schemes have the same \bar{A} because the scaling factor s_i is the same for all the nodes. Hence, ordering on the basis of $|w_i|$ is the same as ordering on the basis of $|y_i|$. As μ increases, the distributions become more separated and \bar{A} decreases.

As ρ increases, two factors counteract each other. First, as ρ increases, the measurements are more correlated. Thus, each additional transmission conveys less information. Hence, more transmissions are needed to make a decision. Second, as ρ increases, the generalized variance decreases. This implies that it becomes easier to distinguish the two distributions since the measurements become more clustered around the mean. When ρ increases from 0 to 0.6, the first factor dominates. As a result, \bar{A} increases. However, when $\rho > 0.6$, the second factor becomes dominant. This leads to the marginal decrease in \bar{A} as ρ increases. For $\mu = 5$, the two distributions are already so separated that the increased clustering around the mean as ρ increases has a negligible impact on the ability of the FN to distinguish between the two distributions. Thus, the first factor dominates and \bar{A} increases as ρ increases.

V. CONCLUSION

We proposed CA-OTS for the binary hypothesis testing problem with correlated Gaussian statistics for the general case in which the mean and covariance of the two hypotheses were different. CA-OTS required far fewer transmissions on average than UTS, in which all nodes transmitted, without compromising on the error probability. For the shift-in-covariance problem, in which the two hypotheses differed only in their covariance matrices, we proposed two novel approaches that brought out the importance of designing the metric and the decision rules. For CovShift-OTS, our upper bound on the average number of transmissions showed that the FN required only one transmission to make a decision at large SNRs. For the shift-in-mean problem, in which only the mean vectors were different, we proposed MeanShift-OTS, which again used a different metric and decision rules. We also derived an upper bound, which showed that as the difference between the mean vectors increased, MeanShift-OTS saved at least half the number of transmissions compared to UTS. Future work includes incorporating power adaptation to improve energy efficiency and modeling deep channel fades, due to which some transmissions may be missed by the FN.

APPENDIX

A. Decision Rules for CA-OTS

In (8), the decision statistic is a sum of four terms. The first two terms, $\mathbf{r}_k^T \mathbf{P}_k \mathbf{r}_k$ and $\mathbf{m}_k^T \mathbf{r}_k$, and thus their sum \mathbf{a}_k , are in terms of measurements that are known to the FN. We bound the third term $\tilde{\mathbf{u}}_{N-k}^T \tilde{\mathbf{T}}_{N-k} \tilde{\mathbf{u}}_{N-k}$ as follows. Since $\tilde{\mathbf{T}}_{N-k}$ is a symmetric matrix, from [32, Ch. A.5.2], we get

$$\begin{aligned} & \lambda_{\min}(\tilde{\mathbf{T}}_{N-k}) \sum_{j \in \mathbb{N} \setminus \{[1], [2], \dots, [k]\}} y_j^2 \\ & \leq \tilde{\mathbf{u}}_{N-k}^T \tilde{\mathbf{T}}_{N-k} \tilde{\mathbf{u}}_{N-k} \\ & \leq \lambda_{\max}(\tilde{\mathbf{T}}_{N-k}) \sum_{j \in \mathbb{N} \setminus \{[1], [2], \dots, [k]\}} y_j^2. \end{aligned} \quad (46)$$

Since $-|y_{[k]}| < y_j < |y_{[k]}|$, for $j \in \mathbb{N} \setminus \{[1], [2], \dots, [k]\}$, we get $0 < \sum_{j \in \mathbb{N} \setminus \{[1], [2], \dots, [k]\}} y_j^2 < (N-k)y_{[k]}^2$. From the definition of $\tilde{\mathbf{v}}_{N-k}$ in Section III-A, we can bound $\tilde{\mathbf{v}}_{N-k}^T \tilde{\mathbf{u}}_{N-k}$ as

$$-\frac{\mathbf{b}_k}{2} < \tilde{\mathbf{v}}_{N-k}^T \tilde{\mathbf{u}}_{N-k} < \frac{\mathbf{b}_k}{2}, \quad (47)$$

where \mathbf{b}_k is defined in the result statement. To bound $\tilde{\mathbf{u}}_{N-k}^T \tilde{\mathbf{T}}_{N-k} \tilde{\mathbf{u}}_{N-k} + 2\tilde{\mathbf{v}}_{N-k}^T \tilde{\mathbf{u}}_{N-k}$, we need to consider the signs of $\lambda_{\min}(\tilde{\mathbf{T}}_{N-k})$ and $\lambda_{\max}(\tilde{\mathbf{T}}_{N-k})$. This gives rise to the following three cases:

1. If $\lambda_{\min}(\tilde{\mathbf{T}}_{N-k}) \geq 0$: In this case, $0 \leq \tilde{\mathbf{u}}_{N-k}^T \tilde{\mathbf{T}}_{N-k} \tilde{\mathbf{u}}_{N-k} < \lambda_{\max}(\tilde{\mathbf{T}}_{N-k})(N-k)y_{[k]}^2$. Substituting this and (47) in (8) yields

$$d(\mathbf{y}) < \mathbf{a}_k + \lambda_{\max}(\tilde{\mathbf{T}}_{N-k})(N-k)y_{[k]}^2 + \mathbf{b}_k, \quad (48)$$

$$d(\mathbf{y}) > \mathbf{a}_k - \mathbf{b}_k. \quad (49)$$

Therefore, if

$$\mathbf{a}_k < \beta_{\text{gen}} - \lambda_{\max}(\tilde{\mathbf{T}}_{N-k})(N-k)y_{[k]}^2 - \mathbf{b}_k, \quad (50)$$

then it follows from (48) that $d(\mathbf{y}) < \beta_{\text{gen}}$. Thus, the FN should decide H_0 . Similarly, if

$$\mathbf{a}_k > \beta_{\text{gen}} + \mathbf{b}_k, \quad (51)$$

then it follows from (49) that $d(\mathbf{y}) > \beta_{\text{gen}}$. Thus, the FN should decide H_1 , as is done in UTS.

2. If $\lambda_{\min}(\tilde{\mathbf{T}}_{N-k}) < 0 < \lambda_{\max}(\tilde{\mathbf{T}}_{N-k})$: Here, $\lambda_{\min}(\tilde{\mathbf{T}}_{N-k})(N-k)y_{[k]}^2 < \tilde{\mathbf{u}}_{N-k}^T \tilde{\mathbf{T}}_{N-k} \tilde{\mathbf{u}}_{N-k} < \lambda_{\max}(\tilde{\mathbf{T}}_{N-k})(N-k)y_{[k]}^2$. The upper bound on $d(\mathbf{y})$ is the same as (48). Hence, it can be shown that the condition for deciding H_0 is the same as that of the first case. This yields (12).

Since $\lambda_{\min}(\tilde{\mathbf{T}}_{N-k})(N-k)y_{[k]}^2 < \tilde{\mathbf{u}}_{N-k}^T \tilde{\mathbf{T}}_{N-k} \tilde{\mathbf{u}}_{N-k}$, it follows that

$$d(\mathbf{y}) > \mathbf{a}_k + \lambda_{\min}(\tilde{\mathbf{T}}_{N-k})(N-k)y_{[k]}^2 - \mathbf{b}_k. \quad (52)$$

Thus, if

$$\mathbf{a}_k > \beta_{\text{gen}} - \lambda_{\min}(\tilde{\mathbf{T}}_{N-k})(N-k)y_{[k]}^2 + \mathbf{b}_k, \quad (53)$$

then, from (52), $d(\mathbf{y}) > \beta_{\text{gen}}$. Hence, the FN should decide H_1 .

3. If $\lambda_{\max}(\tilde{\mathbf{T}}_{N-k}) \leq 0$: In this case, $\lambda_{\min}(\tilde{\mathbf{T}}_{N-k})(N-k)y_{[k]}^2 < \tilde{\mathbf{u}}_{N-k}^T \tilde{\mathbf{T}}_{N-k} \tilde{\mathbf{u}}_{N-k} \leq 0$. Thus,

$$d(\mathbf{y}) < \mathbf{a}_k + \mathbf{b}_k. \quad (54)$$

Therefore, if

$$\mathbf{a}_k < \beta_{\text{gen}} - \mathbf{b}_k, \quad (55)$$

then, from (54), $d(\mathbf{y}) < \beta_{\text{gen}}$ and the FN should decide H_0 . The lower bound on $d(\mathbf{y})$ is the same as (52). Hence, the condition for deciding H_1 is the same as that in the second case.

When $k = N$, the FN has received all the N measurements. Thus, it employs (6) to decide.

B. Decision Rules for CovShift-OTS

Here, we use the following eigenvalue bounds on \mathbf{G} :

$$\lambda_{\min}(\mathbf{G}) \mathbf{z}^T \mathbf{z} \leq d_{\text{cov}}(\mathbf{z}) \leq \lambda_{\max}(\mathbf{G}) \mathbf{z}^T \mathbf{z}. \quad (56)$$

Writing $\mathbf{z}^T \mathbf{z}$ in terms of the mean-shifted measurements received from the nodes $[1], [2], \dots, [k]$, and the ones not yet received from the nodes $[k+1], [k+2], \dots, [N]$, we get

$$\mathbf{z}^T \mathbf{z} = \sum_{i=1}^k z_{[i]}^2 + \sum_{j=k+1}^N z_{[j]}^2. \quad (57)$$

Since the nodes transmit in the decreasing order of the absolute values of their mean-shifted measurements, we have $0 \leq z_{[j]}^2 < z_{[k]}^2$, for $1 \leq k < j \leq N$.

As in Appendix A, we need to consider the following three cases to apply this inequality:

1. If $\lambda_{\min}(\mathbf{G}) \geq 0$: In this case, the above discussions imply the following:

$$\lambda_{\min}(\mathbf{G}) \left(\sum_{i=1}^k z_{[i]}^2 \right) < d_{\text{cov}}(\mathbf{z}), \quad (58)$$

$$d_{\text{cov}}(\mathbf{z}) < \lambda_{\max}(\mathbf{G}) \left(\sum_{i=1}^k z_{[i]}^2 + (N-k)z_{[k]}^2 \right). \quad (59)$$

Therefore, if $\lambda_{\min}(\mathbf{G}) \left(\sum_{i=1}^k z_{[i]}^2 \right) > \beta_{\text{cov}}$, then, from (58), we get $d_{\text{cov}}(\mathbf{z}) > \beta_{\text{cov}}$. Hence, the FN should decide H_1 , just as UTS would after receiving all N measurements. Again, if

$$\lambda_{\max}(\mathbf{G}) \left(\sum_{i=1}^k z_{[i]}^2 \right) < \beta_{\text{cov}} - \lambda_{\max}(\mathbf{G}) (N-k)z_{[k]}^2, \quad (60)$$

then, from (59), $d_{\text{cov}}(\mathbf{z}) < \beta_{\text{cov}}$ and the FN should decide H_0 .

2. If $\lambda_{\min}(\mathbf{G}) < 0 < \lambda_{\max}(\mathbf{G})$: In this case,

$$\lambda_{\min}(\mathbf{G}) \left(\sum_{i=1}^k z_{[i]}^2 + (N-k)z_{[k]}^2 \right) < d_{\text{cov}}(\mathbf{z}). \quad (61)$$

Therefore, if $\lambda_{\min}(\mathbf{G}) \left(\sum_{i=1}^k z_{[i]}^2 + (N-k)z_{[k]}^2 \right) > \beta_{\text{cov}}$, then, from (61), $d_{\text{cov}}(\mathbf{z}) > \beta_{\text{cov}}$ and the FN should decide H_1 . The decision rule for H_0 is obtained using a similar logic.

3. If $\lambda_{\max}(\mathbf{G}) \leq 0$: In this case, we get $d_{\text{cov}}(\mathbf{z}) < \lambda_{\max}(\mathbf{G}) \left(\sum_{i=1}^k z_{[i]}^2 \right)$. Thus, the FN will decide H_0 if $\lambda_{\max}(\mathbf{G}) \left(\sum_{i=1}^k z_{[i]}^2 \right) < \beta_{\text{cov}}$. The decision rule for H_1 follows similarly.

As before, when $k = N$, the FN decides based on (16).

C. CovShift-OTS: Upper Bound on the Average Number of Transmissions

We first consider the case when $\lambda_{\min}(\mathbf{G}) \geq 0$. The average number of transmissions \bar{A} equals

$$\bar{A} = N - \bar{S}, \quad (62)$$

where \bar{S} is the average number of transmissions saved. Let D_1 be the event that the FN decides after receiving one transmission. In this case, the number of transmissions saved is $N - 1$.

Let $\bar{\nu} = \Pr(D_1)$. Clearly, $\bar{S} \geq (N-1)\bar{\nu}$. From the decision rules in (17) and (18), the event D_1 occurs if $\lambda_{\min}(\mathbf{G}) z_{[1]}^2 > \beta_{\text{cov}}$ and the FN decides H_1 , or $\lambda_{\max}(\mathbf{G}) z_{[1]}^2 < \beta_{\text{cov}} - \lambda_{\max}(\mathbf{G}) (N-1) z_{[1]}^2$ and the FN decides H_0 . Substituting the expressions for α_1 and α_0 from (24), $\bar{\nu}$ can be written as

$$\bar{\nu} = \Pr\left(\left[z_{[1]}^2 > \alpha_1 \right] \cup \left[z_{[1]}^2 < \frac{\alpha_0}{N} \right] \right). \quad (63)$$

If $\beta_{\text{cov}} < 0$, then $\alpha_1 \leq \alpha_0 < 0$ and $\bar{\nu} = 1$. Thus, $\bar{S} \geq N - 1$.

Next, consider the case when $\beta_{\text{cov}} \geq 0$. Here, $0 \leq \alpha_0 \leq \alpha_1$ and the events $z_{[1]}^2 > \alpha_1$ and $z_{[1]}^2 < (\alpha_0/N)$ are mutually exclusive. Using this in (63), we get

$$\bar{\nu} = \Pr\left(z_{[1]}^2 > \alpha_1 \right) + \Pr\left(z_{[1]}^2 < \frac{\alpha_0}{N} \right). \quad (64)$$

We now compute the two probability terms in (64) separately.

a) *Computing $\Pr(z_{[1]}^2 > \alpha_1)$* : We know that $\Pr(z_{[1]}^2 > \alpha_1) = 1 - \Pr(z_{[1]}^2 < \alpha_1)$. Conditioning on the hypotheses H_0 and H_1 and applying the law of total probability, we get

$$\Pr(z_{[1]}^2 > \alpha_1) = 1 - \sum_{h=0}^1 \zeta_h \Pr(z_{[1]}^2 < \alpha_1 | H_h). \quad (65)$$

Note that $z_{[1]}^2 < \alpha_1$ implies $z_i^2 < \alpha_1, \forall 1 \leq i \leq N$. Thus,

$$\begin{aligned} & \Pr(z_{[1]}^2 < \alpha_1 | H_h) \\ &= \Pr(-\sqrt{\alpha_1} < z_i < \sqrt{\alpha_1}, \forall 1 \leq i \leq N | H_h). \end{aligned} \quad (66)$$

Under H_0 , the RVs z_i are independent and identically distributed (i.i.d.) Gaussian RVs with zero mean and variance γ_0 . Thus,

$$\begin{aligned} & \Pr(-\sqrt{\alpha_1} < z_i < \sqrt{\alpha_1}, \forall 1 \leq i \leq N | H_0) \\ &= \left[1 - 2Q\left(\sqrt{\frac{\alpha_1}{\gamma_0}} \right) \right]^N = \delta_0(\alpha_1), \end{aligned} \quad (67)$$

where $\delta_0(\cdot)$ is defined in the result statement.

To calculate $\Pr(-\sqrt{\alpha_1} < z_i < \sqrt{\alpha_1}, \forall 1 \leq i \leq N | H_1)$, we use the following trick that holds for the product-correlation model. The RVs z_i can be written as [22]

$$z_i = \sigma_i \left(\sqrt{1 - \theta_i^2} \Psi_i + \theta_i \Psi \right), \quad \text{for } 1 \leq i \leq N, \quad (68)$$

where $\Psi, \Psi_1, \Psi_2, \dots, \Psi_N$ are $(N+1)$ i.i.d. Gaussian RVs with zero mean and unit variance. Let $\xi = \Pr(-\sqrt{\alpha_1} < z_i < \sqrt{\alpha_1}, \forall 1 \leq i \leq N | H_1)$. Substituting z_i from (68), conditioning on Ψ , and applying the law of total probability yields

$$\begin{aligned} \xi &= \mathbb{E}_{\Psi} \left[\Pr \left(\frac{1}{\sqrt{1 - \theta_i^2}} \left[-\sqrt{\frac{\alpha_1}{\sigma_i^2}} - \theta_i \Psi \right] < \Psi_i \right. \right. \\ &< \left. \left. \frac{1}{\sqrt{1 - \theta_i^2}} \left[\sqrt{\frac{\alpha_1}{\sigma_i^2}} - \theta_i \Psi \right], \forall 1 \leq i \leq N \right) \right], \quad (69) \\ &= \mathbb{E}_{\Psi} \left[\prod_{i=1}^N \left[Q \left(\frac{1}{\sqrt{1 - \theta_i^2}} \left[-\sqrt{\frac{\alpha_1}{\sigma_i^2}} - \theta_i \Psi \right] \right) \right. \right. \\ &\left. \left. - Q \left(\frac{1}{\sqrt{1 - \theta_i^2}} \left[\sqrt{\frac{\alpha_1}{\sigma_i^2}} - \theta_i \Psi \right] \right) \right] \right]. \quad (70) \end{aligned}$$

The last step follows because $\Psi_1, \Psi_2, \dots, \Psi_N$, and Ψ are i.i.d. Substituting the PDF of Ψ yields

$$\xi = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \prod_{i=1}^N \left[Q \left(\frac{1}{\sqrt{1-\theta_i^2}} \left[-\sqrt{\frac{\alpha_1}{\sigma_i^2}} - \theta_i \psi \right] \right) - Q \left(\frac{1}{\sqrt{1-\theta_i^2}} \left[\sqrt{\frac{\alpha_1}{\sigma_i^2}} - \theta_i \psi \right] \right) \right] e^{-\frac{\psi^2}{2}} d\psi. \quad (71)$$

From the definition of $\delta_1(\cdot)$ in (29), we get $\xi = \delta_1(\alpha_1)$. Substituting (67) and (71) in (65) yields

$$\Pr(z_{[1]}^2 > \alpha_1) = [1 - \delta_0(\alpha_1)] \zeta_0 + [1 - \delta_1(\alpha_1)] \zeta_1. \quad (72)$$

b) *Computing $\Pr(z_{[1]}^2 < (\alpha_0/N))$* : Conditioning on H_0 and H_1 and applying the law of total probability, we get $\Pr(z_{[1]}^2 < (\alpha_0/N)) = \sum_{h=0}^1 \zeta_h \Pr(z_{[1]}^2 < (\alpha_0/N) | H_h)$. As above,

$$\Pr(z_{[1]}^2 < \frac{\alpha_0}{N}) = \delta_0\left(\frac{\alpha_0}{N}\right) \zeta_0 + \delta_1\left(\frac{\alpha_0}{N}\right) \zeta_1. \quad (73)$$

Adding (72) and (73) yields the expression for $\bar{\nu}$ in (26) when $\beta_{\text{cov}} \geq 0$.

The lower bound on \bar{S} when $\lambda_{\min}(\mathbf{G}) < 0 < \lambda_{\max}(\mathbf{G})$ is obtained in a similar manner by considering the decision rules in (19) and (20). Lastly, the lower bound on \bar{S} when $\lambda_{\max}(\mathbf{G}) \leq 0$ is obtained in a similar manner by considering the decision rules in (21) and (22).

D. MeanShift-OTS: Upper Bound on the Average Number of Transmissions

Let \bar{S}_1 denote the average number of transmissions saved when the event in (41) occurs, which makes the FN decide on H_1 . Similarly, let \bar{S}_0 denote the average number of transmissions saved when the event in (42) occurs, which makes the FN decide on H_0 . Since these events are mutually exclusive, the average number of transmissions saved \bar{S} is given by

$$\bar{S} = \bar{S}_1 + \bar{S}_0. \quad (74)$$

a) *Lower Bound on \bar{S}_1* : Let τ_1 be the number of transmissions after which the FN decides H_1 . Then, proceeding along lines similar to the proof of Theorem 1 in [15], we have

$$\bar{S}_1 = \sum_{k=1}^N (N-k) \Pr(\tau_1 = k) \geq \sum_{k=1}^{\lceil \frac{N}{2} \rceil + 1} (N-k) \Pr(\tau_1 = k). \quad (75)$$

For $1 \leq k \leq \lceil N/2 \rceil + 1$, we have $\lfloor N/2 \rfloor - 1 \leq N-k \leq N-1$. Thus, \bar{S}_1 in (75) can be lower bounded further as

$$\bar{S}_1 \geq \left(\left\lfloor \frac{N}{2} \right\rfloor - 1 \right) \Pr\left(\tau_1 \leq \left\lfloor \frac{N}{2} \right\rfloor + 1\right). \quad (76)$$

When the FN decides H_1 in at most $\lceil N/2 \rceil + 1$ transmissions, it follows from (41) that $\sum_{i=1}^{\lceil N/2 \rceil + 1} w_{[i]} > \beta_{\text{mean}} +$

$(\lfloor N/2 \rfloor - 1) |w_{\lceil N/2 \rceil + 1}|$. Hence,

$$\begin{aligned} \Pr\left(\tau_1 \leq \left\lfloor \frac{N}{2} \right\rfloor + 1\right) \\ = \Pr\left(\sum_{i=1}^{\lceil \frac{N}{2} \rceil + 1} w_{[i]} > \beta_{\text{mean}} + \left(\left\lfloor \frac{N}{2} \right\rfloor - 1\right) |w_{\lceil \frac{N}{2} \rceil + 1}| \right). \end{aligned} \quad (77)$$

Next we show that $w_i > \beta_{\text{mean}}, \forall 1 \leq i \leq N$, implies $\sum_{i=1}^{\lceil N/2 \rceil + 1} w_{[i]} > \beta_{\text{mean}} + (\lfloor N/2 \rfloor - 1) |w_{\lceil N/2 \rceil + 1}|$. For this, notice that $\beta_{\text{mean}} = (\boldsymbol{\mu}^{(1)})^T \mathbf{R}^{-1} \boldsymbol{\mu}^{(1)} > 0$ since \mathbf{R} is a positive-definite matrix. Hence, $w_i > \beta_{\text{mean}}$ implies that $|w_i| > \beta_{\text{mean}}$. This, in turn, implies that $\sum_{i=1}^{\lceil N/2 \rceil + 1} w_{[i]} > \beta_{\text{mean}} + (\lfloor N/2 \rfloor - 1) |w_{\lceil N/2 \rceil + 1}|$. Hence, from (77), we get

$$\Pr\left(\tau_1 \leq \left\lfloor \frac{N}{2} \right\rfloor + 1\right) \geq \Pr(w_i > \beta_{\text{mean}}, \forall 1 \leq i \leq N). \quad (78)$$

Conditioning on the hypotheses, $\Pr(w_i > \beta_{\text{mean}}, \forall 1 \leq i \leq N)$ can be written as $\Pr(w_i > \beta_{\text{mean}}, \forall 1 \leq i \leq N) = \sum_{h=0}^1 \zeta_h \Pr(w_i > \beta_{\text{mean}}, \forall 1 \leq i \leq N | H_h)$.

Since $w_i = s_i y_i$ and s_i is a constant, it follows that w_1, w_2, \dots, w_N also follow the product-correlation model when conditioned on H_h . Their correlation-coefficient χ_{ij} is given by $\chi_{ij} = \theta_i \theta_j$, if $s_i s_j > 0$, and $\chi_{ij} = -\theta_i \theta_j$, otherwise. Thus, χ_{ij} can be compactly written as $\chi_{ij} = \text{sgn}(s_i s_j) \theta_i \theta_j$. Furthermore, w_i is a Gaussian RV with mean $s_i \mu_i^{(h)}$ under hypothesis $h \in \{0, 1\}$, and variance $s_i^2 \sigma_i^2$. Hence, along similar lines as (68), for $1 \leq i \leq N$, w_i can be expressed in terms of $(N+1)$ i.i.d. Gaussian RVs $\Psi, \Psi_1, \Psi_2, \dots, \Psi_N$ with zero mean and unit variance as

$$w_i = s_i \mu_i^{(h)} + |s_i| \sigma_i \left(\sqrt{1-\theta_i^2} \Psi_i + \text{sgn}(s_i) \theta_i \Psi \right). \quad (79)$$

Let $\Upsilon_h = \Pr(w_i > \beta_{\text{mean}}, \forall 1 \leq i \leq N | H_h)$, for $h \in \{0, 1\}$. Substituting (79) in Υ_h , we get

$$\begin{aligned} \Upsilon_h = \Pr\left(\Psi_i > \frac{1}{\sqrt{1-\theta_i^2}} \left[\frac{\beta_{\text{mean}} - s_i \mu_i^{(h)}}{|s_i| \sigma_i} - \text{sgn}(s_i) \theta_i \Psi \right], \right. \\ \left. \forall 1 \leq i \leq N \right). \end{aligned} \quad (80)$$

Conditioning on Ψ and proceeding along similar lines as (69) and (70) yields

$$\begin{aligned} \Upsilon_h = \\ \mathbb{E}_{\Psi} \left[\prod_{i=1}^N \left[Q \left(\frac{1}{\sqrt{1-\theta_i^2}} \left[\frac{\beta_{\text{mean}} - s_i \mu_i^{(h)}}{|s_i| \sigma_i} - \text{sgn}(s_i) \theta_i \Psi \right] \right) \right] \right]. \end{aligned} \quad (81)$$

Substituting the Gaussian PDF of Ψ and writing the expectation as an integral yields the expression for η_h in (44). Thus, $\Upsilon_h = \eta_h$ and $\Pr(w_i > \beta_{\text{mean}}, \forall 1 \leq i \leq N) = \sum_{h=0}^1 \zeta_h \eta_h$. Substituting this in (78) and then in (76), we get

$$\bar{S}_1 \geq \left(\left\lfloor \frac{N}{2} \right\rfloor - 1 \right) \sum_{h=0}^1 \zeta_h \eta_h. \quad (82)$$

b) *Lower Bound on \bar{S}_0* : Along similar lines as above, \bar{S}_0 can be lower bounded as

$$\bar{S}_0 \geq \left(\left\lfloor \frac{N}{2} \right\rfloor - 1 \right) \sum_{h=0}^1 \zeta_h \kappa_h, \quad (83)$$

where κ_h is defined in (45). Substituting (82) and (83) in (74) and then in (62) yields (43).

ACKNOWLEDGMENT

The authors wish to thank C. R. Murthy, A. Bodas, P. S. Aditya, and R. K. Poricha for discussions on this topic.

REFERENCES

- [1] S. S. Gupta and N. B. Mehta, "Correlation-aware ordered transmissions scheme for energy-efficient detection," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2020, pp. 1–6.
- [2] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *IEEE Commun. Mag.*, vol. 40, no. 8, pp. 102–114, Aug. 2002.
- [3] H. Yetgin, K. T. K. Cheung, M. El-Hajjar, and L. Hanzo, "A survey of network lifetime maximization techniques in wireless sensor networks," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 2, pp. 828–854, 2nd Quart., 2017.
- [4] A. A. Saucan and M. Z. Win, "Information-seeking sensor selection for Ocean-of-Things," *IEEE Internet Things J.*, vol. 7, no. 10, pp. 10072–10088, Oct. 2020.
- [5] W. P. Tay, J. N. Tsitsiklis, and M. Z. Win, "Asymptotic performance of a censoring sensor network," *IEEE Trans. Inf. Theory*, vol. 53, no. 11, pp. 4191–4209, Nov. 2007.
- [6] R. S. Blum and B. M. Sadler, "Energy efficient signal detection in sensor networks using ordered transmissions," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 3229–3235, Jul. 2008.
- [7] Z. N. Rawas, Q. He, and R. S. Blum, "Energy-efficient noncoherent signal detection for networked sensors using ordered transmissions," in *Proc. 45th Annu. Conf. Inf. Sci. Syst.*, Mar. 2011, pp. 1–5.
- [8] K. Cohen and A. Leshem, "Energy-efficient detection in wireless sensor networks using likelihood ratio and channel state information," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 8, pp. 1671–1683, Sep. 2011.
- [9] P. Braca, S. Marano, and V. Matta, "Single-transmission distributed detection via order statistics," *IEEE Trans. Signal Process.*, vol. 60, no. 4, pp. 2042–2048, Apr. 2012.
- [10] L. Hesham, A. Sultan, M. Nafie, and F. Digham, "Distributed spectrum sensing with sequential ordered transmissions to a cognitive fusion center," *IEEE Trans. Signal Process.*, vol. 60, no. 5, pp. 2524–2538, May 2012.
- [11] S. Marano, V. Matta, and P. Willett, "Nearest-neighbor distributed learning by ordered transmissions," *IEEE Trans. Signal Process.*, vol. 61, no. 21, pp. 5217–5230, Nov. 2013.
- [12] Y. Chen, B. M. Sadler, and R. S. Blum, "Ordered transmission for efficient wireless autonomy," in *Proc. 52nd Asilomar Conf. Signals, Syst., Comput.*, Oct. 2018, pp. 1299–1303.
- [13] N. Sriranga, K. G. Nagananda, and R. S. Blum, "Shared channel ordered transmissions for energy-efficient distributed signal detection," *IEEE Commun. Lett.*, vol. 23, no. 1, pp. 96–99, Jan. 2019.
- [14] S. Sen Gupta, S. K. Pallapothu, and N. B. Mehta, "Ordered transmissions for energy-efficient detection in energy harvesting wireless sensor networks," *IEEE Trans. Commun.*, vol. 68, no. 4, pp. 2525–2537, Apr. 2020.
- [15] J. Zhang, Z. Chen, R. S. Blum, X. Lu, and W. Xu, "Ordering for reduced transmission energy detection in sensor networks testing a shift in the mean of a Gaussian graphical model," *IEEE Trans. Signal Process.*, vol. 65, no. 8, pp. 2178–2189, Apr. 2017.
- [16] Y. Chen, R. S. Blum, B. M. Sadler, and J. Zhang, "Testing the structure of a Gaussian graphical model with reduced transmissions in a distributed setting," *IEEE Trans. Signal Process.*, vol. 67, no. 20, pp. 5391–5401, Oct. 2019.
- [17] V. Shah, N. Mehta, and R. Yim, "Optimal timer based selection schemes," *IEEE Trans. Commun.*, vol. 58, no. 6, pp. 1814–1823, Jun. 2010.
- [18] H. Chen, B. Chen, and P. K. Varshney, "A new framework for distributed detection with conditionally dependent observations," *IEEE Trans. Signal Process.*, vol. 60, no. 3, pp. 1409–1419, Mar. 2012.
- [19] A. Jindal and K. Psounis, "Modeling spatially correlated data in sensor networks," *ACM Trans. Sensor Netw.*, vol. 2, no. 4, pp. 466–499, Nov. 2006.
- [20] R. Mallik, "The uniform correlation matrix and its application to diversity," *IEEE Trans. Wireless Commun.*, vol. 6, no. 5, pp. 1619–1625, May 2007.
- [21] Y. Chen and C. Tellambura, "Distribution functions of selection combiner output in equally correlated Rayleigh, Rician, and Nakagami-m fading channels," *IEEE Trans. Commun.*, vol. 52, no. 11, pp. 1948–1956, Nov. 2004.
- [22] N. C. Beaulieu and K. T. Hemachandra, "Novel simple representations for Gaussian class multivariate distributions with generalized correlation," *IEEE Trans. Inf. Theory*, vol. 57, no. 12, pp. 8072–8083, Dec. 2011.
- [23] M. F. A. Ahmed, T. Y. Al-Naffouri, M.-S. Alouini, and G. Turkiyyah, "The effect of correlated observations on the performance of distributed estimation," *IEEE Trans. Signal Process.*, vol. 61, no. 24, pp. 6264–6275, Dec. 2013.
- [24] K. Bai, H. Senol, and C. Tepedelenlioglu, "Outage scaling laws and diversity for distributed estimation over parallel fading channels," *IEEE Trans. Signal Process.*, vol. 57, no. 8, pp. 3182–3192, Aug. 2009.
- [25] A. S. Leong, S. Dey, G. N. Nair, and P. Sharma, "Power allocation for outage minimization in state estimation over fading channels," *IEEE Trans. Signal Process.*, vol. 59, no. 7, pp. 3382–3397, Jul. 2011.
- [26] R. K. Mallik, "The exponential correlation matrix: Eigen-analysis and applications," *IEEE Trans. Wireless Commun.*, vol. 17, no. 7, pp. 4690–4705, Jul. 2018.
- [27] L. Scharf and B. Van Veen, "Low rank detectors for Gaussian random vectors," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 35, no. 11, pp. 1579–1582, Nov. 1987.
- [28] S. Zhang, P. Khanduri, and P. K. Varshney, "Distributed sequential detection: Dependent observations and imperfect communication," *IEEE Trans. Signal Process.*, vol. 68, pp. 830–842, Nov. 2020.
- [29] H. V. Poor, *An Introduction to Signal Detection and Estimation*, 2nd ed. New York, NY, USA: Springer-Verlag, 1994.
- [30] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions With Formulas, Graphs, and Mathematical Tables*, 9th ed. New York, NY, USA: Dover, 1972.
- [31] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*, 3rd ed. Hoboken, NJ, USA: Wiley, 2003.
- [32] S. Boyd and L. Vandenberghe, *Convex Optimization*, 7th ed. Cambridge, U.K.: Cambridge Univ. Press, 2009.



Sayan Sen Gupta (Graduate Student Member, IEEE) received the B.Tech. degree in electronics and communications from the Maulana Abul Kalam Azad University of Technology, West Bengal, Kolkata, India, in 2011, and the M.E. degree in electronics and telecommunications from Jadavpur University, Kolkata, India, in 2013. He was with Tata Consultancy Services Ltd., from 2013 to 2015. He is currently pursuing the Ph.D. degree with the Department of Electrical Communication Engineering, Indian Institute of Science, Bengaluru. His

research interests include design and analysis of energy-efficient schemes for wireless sensor networks.



Neelesh B. Mehta (Fellow, IEEE) received the B.Tech. degree in electronics and communications engineering from the IIT Madras in 1996 and the M.S. and Ph.D. degrees in electrical engineering from the California Institute of Technology, Pasadena, USA, in 1997 and 2001, respectively. He is currently a Professor with the Department of Electrical Communication Engineering, Indian Institute of Science, Bengaluru. He is a fellow of the Indian National Science Academy, the Indian National Academy of Engineering, and the National Academy of Sciences India. He was a recipient of the Shanti Swarup Bhatnagar Award, the Khosla Award, the Vikram Sarabhai Research Award, and the Swarnjayanti Fellowship. He served on the Board of Governors for the IEEE ComSoc from 2012 to 2015. He served on the Executive Editorial Committee for IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS from 2014 to 2017, and served as its Chair from 2017 to 2018. He has served as an Editor for IEEE TRANSACTIONS ON COMMUNICATIONS and IEEE WIRELESS COMMUNICATION LETTERS in the past.