Research Article

Censored Distributed Space-Time Coding for Wireless Sensor Networks

S. Yiu and R. Schober

Department of Electrical and Computer Engineering, The University of British Columbia, 2356 Main Mall, Vancouver, BC, Canada V6T 1Z4

Correspondence should be addressed to S. Yiu, simony@ece.ubc.ca

We consider the application of distributed space-time coding in wireless sensor networks (WSNs). In particular, sensors use a common noncoherent distributed space-time block code (DSTBC) to forward their local decisions to the fusion center (FC) which makes the final decision. We show that the performance of distributed space-time coding is negatively affected by erroneous sensor decisions caused by observation noise. To overcome this problem of error propagation, we introduce censored distributed space-time coding where only reliable decisions are forwarded to the FC. The optimum noncoherent maximum-likelihood and a low-complexity, suboptimum generalized likelihood ratio test (GLRT) FC decision rules are derived and the performance of the GLRT decision rule is analyzed. Based on this performance analysis we derive a gradient algorithm for optimization of the local decision/censoring threshold. Numerical and simulation results show the effectiveness of the proposed censoring scheme making distributed space-time coding a prime candidate for signaling in WSNs.

1. INTRODUCTION

In recent years, wireless sensor networks (WSNs) have been gaining popularity in a wide range of military and civilian applications such as environmental monitoring, health care, and control. A typical WSN consists of a number of geographically distributed sensors and a fusion center (FC). The low-cost and low-power sensors make local observations of the hypotheses under test and communicate with the FC. *Centralized* detection schemes require the sensors to transmit their real-valued observations to the FC. However, this automatically translates into the unrealistic assumption of an infinite-bandwidth communication channel. In reality, the WSN has to work in a bandlimited environment. Moreover, as communication is a key energy consumer in a WSN, it is desirable to process the observation data as much as possible at the local sensors to reduce the number of bits that have to be transmitted over the communication channel. Therefore, the sensors typically make local decisions which are then transmitted to the FC where the final decision is made [1–5].

The resulting *decentralized* detection problem has a long and rich history. The decentralized optimum hypothesis testing problem was first formulated in [1] to provide a theoretical framework for detection with distributed sensors. Traditionally, the local decisions are assumed to be transmitted to the FC through perfect, error-free channels [1–6]. Realistically, the sensors typically work in harsh environments and therefore, fading and noise should be taken into account.

The problem of fusing sensor decisions over noisy and fading channels was considered in [7, 8]. The fusion rules developed in [7] require instantaneous channel-state information (CSI). While the fusion rules in [8] do not require amplitude CSI, they still assume perfect phase estimation/synchronization. However, obtaining any form of CSI may not be feasible in large-scale WSNs and cheap sensors make phase synchronization challenging. To avoid these problems, simple ON/OFF keying and corresponding fusion rules were considered in [9]. Furthermore, power efficiency is improved in [9] by employing a simple form of censoring [10], where the sensors transmit only reliable decisions to the FC. The schemes in [7–9] assume orthogonal channels

between the sensors and the FC, which entail a large required bandwidth especially in dense WSNs with a large number of sensors.

To overcome the bandwidth limitations of orthogonal transmission in WSNs, the application of coherent distributed space-time coding was proposed in [11]. In particular, in [11] each sensor is randomly assigned a column of Alamouti's space-time block code (STBC) [12] and it is assumed that only two sensors are active randomly at any time. The quantized observations are encoded by the sensors using the respective preassigned columns of the STBC and transmitted to the FC via a common, noorthogonal channel. Since there are typically more sensors than STBC columns, the same column has to be assigned to more than one sensor resulting in a diversity order of 1. The performance degradation due to the diversity loss and the observation noise is analyzed in [11].

We point out that distributed space-time coding is usually employed in relay networks where a cyclic redundancy check (CRC) code can be used to avoid the retransmission of incorrect decisions by the relays [13–15]. In this context, selection relaying first introduced in [16] has some similarities to censoring in sensor networks [9, 10]. However, while in selection relaying the decision whether a relay retransmits a packet or not depends on the instantaneous CSI of the source-relay channel, the censoring decision depends on the observation noise at the sensor. Furthermore, relaying decisions in selection relaying are made on a packet-by-packet basis enabling coherent detection at the destination node but censor decisions are performed on a symbol-by-symbol basis making coherent data fusion at the FC practically impossible.

In this paper, we consider noncoherent distributed space-time block coding for transmission of censored sensor decisions in WSNs. In particular, we make the following contributions.

(i) We show that the noncoherent distributed STBCs (DSTBCs) introduced in [14] eliminate the various restrictions and drawbacks of the coherent scheme in [11].

(ii) Moreover, it is shown that censoring of local decisions is essential for the efficient application of distributed space-time coding in WSNs.

(iii) We derive the optimum maximum-likelihood (ML) and a suboptimum generalized likelihood ratio test (GLRT) noncoherent FC decision rules for the proposed signaling scheme.

(iv) The bit-error rate (BER) at the FC for the GLRT decision rule is characterized analytically.

(v) Based on the analytical expression for the BER, we devise a gradient algorithm for calculation of the optimum local decision/censoring threshold.

(vi) Our numerical and simulation results show the effectiveness of the proposed transmission scheme and the ability of the noncoherent DSTBC to achieve a diversity gain in WSNs.

This paper is organized as follows. In Section 2, we present the system model and introduce the proposed transmission scheme for WSNs. In Section 3, we derive the
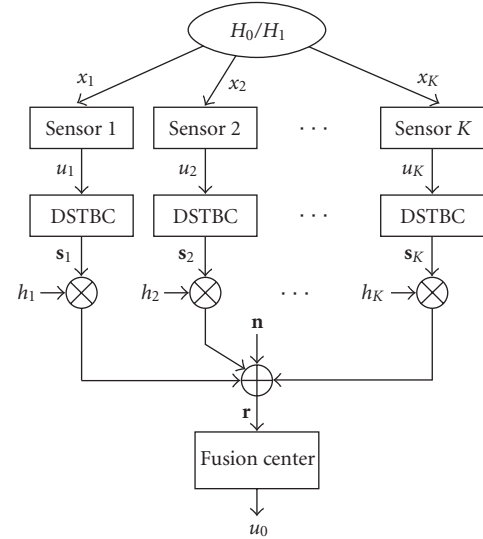

FIGURE 1: Parallel fusion model with K sensors and one FC. A censored DSTBC is used for transmission from the sensors to the FC.

ML and GLRT noncoherent FC decision rules and analyze the performance of the GLRT decision rule. A gradient algorithm for optimization of the local decision/censoring threshold is provided in Section 4. Simulation and numerical results are given in Section 5, while some conclusions are drawn in Section 6.

Notation. In this paper, bold upper case and lower case letters denote matrices and vectors, respectively. $[\cdot]^T$, $[\cdot]^H$, $\varepsilon\{\cdot\}$, $||\cdot||_2$, $|\cdot|$, and $\cup$ denote transposition, Hermitian transposition, statistical expectation, the L_2-norm of a vector, the cardinality of a set, and the union of two sets, respectively. In addition, $Q(x) \triangleq 1/\sqrt{2\pi}\int_x^\infty e^{-t^2/2}\,dt$, $\mathbf{I}_X$, $\mathbf{0}_{X\times Y}$, and $j \triangleq \sqrt{-1}$ denote the Gaussian Q-function, the $X\times X$ identity matrix, the $X \times Y$ all zeros matrix, and the imaginary unit, respectively.

2. SYSTEM MODEL

The binary hypothesis testing problem under consideration is illustrated in Figure 1, where a set $\mathcal{K} \triangleq \{1, 2, \ldots, K\}$ of K distributed sensors tries to determine the true state of nature H as being H_0 (the null hypothesis) or H_1 (target-present hypothesis). Typical applications for binary hypothesis testing include seismic detection, forest fire detection, and environmental monitoring. The a priori probabilities of the two hypotheses H_0 and H_1 are denoted as $P(H_0)$ and $P(H_1)$, respectively. We assume that $P(H_0) = P(H_1) = 0.5$ throughout this paper. The details of the system model will be discussed in the following subsections.

2.1. Local sensor decisions

We assume that the sensor observations are described by

$$
\begin{aligned}
H_0 &: x_k = -1 + n_k, \quad k \in \mathcal{K}, \\
H_1 &: x_k = 1 + n_k, \quad k \in \mathcal{K},
\end{aligned}
\tag{1}
$$

where the local observation noise samples n_k, $k \in \mathcal{K}$, are independent and identically distributed (i.i.d.). For convenience and similar to [8, 9, 11], we assume identical sensors in this paper and model n_k as real-valued additive white Gaussian noise (AWGN) with zero mean and variance $\sigma^2 \triangleq \varepsilon\{n_k^2\}$, $k \in \mathcal{K}$. We note, however, that the generalization of our results to nonidentical sensors (e.g., sensors with different noise variances) is also possible.

Upon receiving its own observation, each sensor makes a ternary local decision:

$$ u_k = \begin{cases} -1 & \text{if } x_k < -d, \\ 1 & \text{if } x_k > d, \qquad k \in \mathcal{K}, \\ 0 & \text{otherwise}, \end{cases} \qquad (2) $$

where d is the nonnegative decision/censoring threshold. While $u_k = -1$ and $u_k = 1$ correspond to hypotheses H_0 and H_1, respectively, $u_k = 0$ corresponds to a decision that is deemed unreliable by the sensor and thus censored. For future reference, we denote the sets of sensors with $u_k = 0$, $u_k = -1$, and $u_k = 1$ by $\mathscr{S}$, $\mathcal{H}_0$, and $\mathcal{H}_1$, respectively. Note that $\mathcal{K} = \mathscr{S} \cup \mathcal{H}_0 \cup \mathcal{H}_1$.

It is not difficult to show that the probabilities of correct and wrong sensor decision are given by

$$ \begin{aligned} P_c &= Q\left(\frac{d-1}{\sigma}\right), \\ P_w &= Q\left(\frac{d+1}{\sigma}\right), \end{aligned} \qquad (3) $$

respectively. The probability that a decision is censored is given by

$$ P_s = 1 - P_c - P_w = 1 - Q\left(\frac{d-1}{\sigma}\right) - Q\left(\frac{d+1}{\sigma}\right). \qquad (4) $$

2.2. Noncoherent distributed space-time coding

The general concept of DSTBC was originally proposed in [13] to achieve a diversity gain in cooperative networks with decode-and-forward relaying. The DSTBC scheme in [14] is particularly attractive for application in networks with a large number of nodes since its decoding complexity is independent of the total number of nodes. This scheme consists of a code $\mathcal{C}$ and a set of signature vectors $\mathcal{G}$. The active relay nodes[1] encode the (correctly decoded) source information using a $T \times N$ code matrix $\boldsymbol{\Phi} \in \mathcal{C}$. Each active relay transmits a linear combination of the columns of the information-carrying matrix $\boldsymbol{\Phi}$. The linear combination coefficients for each node are unique and are collected in a signature vector $\mathbf{g}_k \in \mathcal{G}$, $\|\mathbf{g}_k\|_2^2 = 1$, $k \in \mathcal{K}$, of length N.

In this work, we consider the application of the DSTBC scheme in [14] in WSNs. In particular, sensors encode their local decisions using a noncoherent DSTBC. Since we consider here a binary hypothesis testing problem, $\mathcal{C} = \{\boldsymbol{\Phi}_0, \boldsymbol{\Phi}_1\}$

has only two elements. To optimize performance under noncoherent detection, we choose $\boldsymbol{\Phi}_0$ and $\boldsymbol{\Phi}_1$ to be orthogonal, that is, $\boldsymbol{\Phi}_0^H \boldsymbol{\Phi}_1 = \mathbf{0}_{N \times N}$ and $\boldsymbol{\Phi}_\nu^H \boldsymbol{\Phi}_\nu = \mathbf{I}_N$, $\nu \in \{0, 1\}$ (cf. [17]). Each sensor is assigned a unique signature vector $\mathbf{g}_k \in \mathcal{G}$, $\|\mathbf{g}_k\|_2^2 = 1$, $k \in \mathcal{K}$, of length N. For the design of deterministic and random signature vector sets $\mathcal{G}$, we refer to [14, 15] , respectively. The transmitted signal of sensor k is given by

$$ \mathbf{s}_k = \begin{cases} \sqrt{E}\boldsymbol{\Phi}_0\mathbf{g}_k & \text{if } k \in \mathcal{H}_0, \\ \sqrt{E}\boldsymbol{\Phi}_1\mathbf{g}_k & \text{if } k \in \mathcal{H}_1, \\ \mathbf{0}_{T \times 1} & \text{if } k \in \mathscr{S}, \end{cases} \qquad (5) $$

where E denotes the transmitted energy of sensor k per codeword. We note that sensor k transmits the T elements of $\mathbf{s}_k$ in T consecutive symbol intervals. The total average transmitted energy per information bit is given by $E_b = EK(P_w + P_c)$.

2.3. Channel model

We assume that the sensors transmit time synchronously and that the sensor-FC channels are frequency-nonselective and time-invariant for at least T symbol intervals.[2] Therefore, using the equivalent complex baseband representation of bandpass signals, the signal samples received at the FC in T consecutive symbol intervals can be expressed as

$$ \mathbf{r} = \sum_{k \in \mathcal{H}_0 \cup \mathcal{H}_1} \mathbf{s}_k h_k + \mathbf{n} = \sqrt{E}\boldsymbol{\Phi}_0\mathbf{G}_{\mathcal{H}_0}\mathbf{h}_{\mathcal{H}_0} + \sqrt{E}\boldsymbol{\Phi}_1\mathbf{G}_{\mathcal{H}_1}\mathbf{h}_{\mathcal{H}_1} + \mathbf{n}, \qquad (6) $$

where h_k and $\mathbf{n}$ denote the fading gain of sensor k and a complex AWGN vector, respectively. The columns of the $N \times |\mathcal{H}_0|$ matrix $\mathbf{G}_{\mathcal{H}_0}$ and $N \times |\mathcal{H}_1|$ matrix $\mathbf{G}_{\mathcal{H}_1}$ contain the signature vectors of the sensors in $\mathcal{H}_0$ and $\mathcal{H}_1$, respectively. The corresponding fading gains are collected in column vectors $\mathbf{h}_{\mathcal{H}_0}$ and $\mathbf{h}_{\mathcal{H}_1}$ which have lengths $|\mathcal{H}_0|$ and $|\mathcal{H}_1|$, respectively. We model the channel gains h_k, $k \in \mathcal{K}$, as i.i.d. zero-mean complex Gaussian random variables (Rayleigh fading) with variance $\sigma_h^2 = \varepsilon\{|h_k|^2\} = 1$.[3] The elements of the noise vector $\mathbf{n}$ have variance $\sigma_n^2 = N_0$, where N_0 denotes the power spectral density of the underlying continuous-time passband noise process.

Equation (6) clearly shows the importance of censoring when applying DSTBCs in WSNs, since incorrect sensor decisions lead to interference. For example, for $H = H_0$, ideally the term involving $\boldsymbol{\Phi}_1$ in (6) would be absent. However, incorrect decisions may cause some sensors to transmit $\sqrt{E}\boldsymbol{\Phi}_1\mathbf{g}_k$ instead of $\sqrt{E}\boldsymbol{\Phi}_0\mathbf{g}_k$. The considered censoring

---

[1] The relays which fail to decode the source packet correctly remain silent.

[2] Time synchronous transmission can be accomplished if the relative delays between the relay nodes are much smaller than the symbol duration. This is usually a reasonable assumption for low-rate WSN applications. We refer the interested reader to [18] for a more detailed discussion on time synchronism in the context of WSNs.

[3] This model is justified if the distance between any pair of sensors is much smaller than the distances between the sensors and the FC. The effect of unequal channel variances is considered in Section 5 (cf. Figure 7).

scheme reduces the number of incorrect decisions (by choosing $d > 0$) at the expense of reducing the number of sensors that make a correct decision. However, this disadvantage is outweighed by the reduction of interference as long as d is not too large (cf. Section 5). We note that censoring was not considered in any of the related publications, for example, [11, 13–15]. For example, in [13–15], DSTBCs were mainly applied for relay purposes, where a CRC code can be used to avoid the retransmission of incorrect decisions.

2.4. *Processing at fusion center (FC)*

The FC makes a decision based on the received vector $\mathbf{r}$ and outputs $u_0 = 1$ if it decides in favor of H_1, and $u_0 = -1$ otherwise. Different decision rules may be applied at the FC differing in performance and complexity. In this context, we note that coherent detection is not feasible in large-scale WSNs since the FC would have to estimate and track the channel gains of all sensors. While (6) suggests that only the effective channels $\sqrt{E}\mathbf{G}_{\mathcal{H}_0}\mathbf{h}_{H_0}$ and $\sqrt{E}\mathbf{G}_{\mathcal{H}_1}\mathbf{h}_{\mathcal{H}_1}$ have to be estimated if distributed space-time coding is applied, this is also not feasible since the sets $\mathcal{H}_0$ and $\mathcal{H}_1$ typically change after T symbol intervals (i.e., for every new sensor decision). Therefore, only noncoherent decision rules will be considered in the next section.

3. FC DECISION RULES AND PERFORMANCE ANALYSIS

In this section, we present the optimum ML and the generalized-likelihood ratio test (GLRT) noncoherent decision rules. In addition, we provide a performance analysis for the GLRT decision rule.

3.1. *Optimum maximum-likelihood (ML) decision rule*

We first provide the optimum ML decision rule. For this purpose, we introduce the likelihood ratio (LR):

$$\Lambda_o(\mathbf{r}) \triangleq \frac{f(\mathbf{r} \mid H_1)}{f(\mathbf{r} \mid H_0)}$$
$$= \frac{\sum_{\mathcal{H}_0, \mathcal{H}_1} f(\mathbf{r} \mid \mathcal{H}_0, \mathcal{H}_1) P(\mathcal{H}_0, \mathcal{H}_1 \mid H_1)}{\sum_{\mathcal{H}_0, \mathcal{H}_1} f(\mathbf{r} \mid \mathcal{H}_0, \mathcal{H}_1) P(\mathcal{H}_0, \mathcal{H}_1 \mid H_0)}, \quad (7)$$

where $P(H_0, \mathcal{H}_1 \mid H_0) = P_c^{|\mathcal{H}_0|} P_s^{|\mathcal{S}|} P_w^{|\mathcal{H}_1|}$ and $P(\mathcal{H}_0, \mathcal{H}_1 \mid H_1) = P_c^{|\mathcal{H}_1|} P_s^{|\mathcal{S}|} P_w^{|\mathcal{H}_0|}$ denote the probabilities that the sets $\mathcal{H}_0, \mathcal{H}_1$ occur for H_0 and H_1, respectively. Since $\mathbf{r}$ conditioned on $\mathcal{H}_0, \mathcal{H}_1$ is a Gaussian vector, the conditional probability density function (pdf) $f(\mathbf{r} \mid \mathcal{H}_0, \mathcal{H}_1)$ is given by

$$f(\mathbf{r} \mid \mathcal{H}_0, \mathcal{H}_1) = \frac{\exp(-\mathbf{r}^H \mathbf{B}\mathbf{r})}{\pi^T \det(\mathbf{B})}, \quad (8)$$

where the $T \times T$ correlation matrix $\mathbf{B}$ is defined as $\mathbf{B} \triangleq \varepsilon\{\mathbf{r}\mathbf{r}^H \mid \mathcal{H}_0, \mathcal{H}_1\} = E(\mathbf{\Phi}_0\mathbf{G}_{\mathcal{H}_0}\mathbf{G}_{\mathcal{H}_0}^H\mathbf{\Phi}_0^H + \mathbf{\Phi}_1\mathbf{G}_{\mathcal{H}_1}\mathbf{G}_{\mathcal{H}_1}^H\mathbf{\Phi}_1^H) + \sigma_n^2\mathbf{I}_T$. Now we can express the ML decision rule at the FC as

$$u_0 = \begin{cases} 1 & \text{if } \Lambda_o(\mathbf{r}) \geq 1, \\ -1 & \text{if } \Lambda_o(\mathbf{r}) < 1. \end{cases} \quad (9)$$

We note that the sums in the numerator and denominator of (7) both have 3^K terms, that is, the complexity of the ML decision rule is of orde $O(3^K)$ and grows exponentially with K. In addition, (8) reveals that for the ML decision rule the FC requires knowledge of the signature vectors of all sensors. These two assumptions make the implementation of the ML decision rule difficult, if not impossible in practice. Therefore, we will provide a low-complexity suboptimum FC decision rule in the next subsection.

3.2. *GLRT decision rule*

The received vector can be expressed as

$$\mathbf{r} = \mathbf{\Phi}\mathbf{h}_{\text{eff}} + \mathbf{n}_{\text{eff}}, \quad \mathbf{\Phi} \in \{\mathbf{\Phi}_0, \mathbf{\Phi}_1\}. \quad (10)$$

If H_0 is the true hypothesis $\mathbf{\Phi} = \mathbf{\Phi}_0$, $\mathbf{h}_{\text{eff}} \triangleq \sqrt{E}\mathbf{G}_{\mathcal{H}_0}\mathbf{h}_{H_0}$, and $\mathbf{n}_{\text{eff}} \triangleq \sqrt{E}\mathbf{\Phi}_1\mathbf{G}_{\mathcal{H}_1}\mathbf{h}_{\mathcal{H}_1} + \mathbf{n}$, while if H_1 is true $\mathbf{\Phi} = \mathbf{\Phi}_1$, $\mathbf{h}_{\text{eff}} \triangleq \sqrt{E}\mathbf{G}_{\mathcal{H}_1}\mathbf{h}_{\mathcal{H}_1}$, and $\mathbf{n}_{\text{eff}} \triangleq \sqrt{E}\mathbf{\Phi}_0\mathbf{G}_{\mathcal{H}_0}\mathbf{h}_{H_0} + \mathbf{n}$.

Equation (10) suggests a two-step GLRT approach for the estimation of the transmitted codewor $\mathbf{\Phi}$. In the first step, $\mathbf{h}_{\text{eff}}$ is estimated assuming $\mathbf{\Phi}$ is known, and in the second step the channel estimate $\hat{\mathbf{h}}_{\text{eff}}$ is used to detect $\mathbf{\Phi}$. Since the correlation matrix of the effective noise $\mathbf{n}_{\text{eff}}$ depends on $\mathbf{G}_{\mathcal{H}_1}$ or $\mathbf{G}_{\mathcal{H}_0}$, the ML estimate for $\mathbf{h}_{\text{eff}}$ and thus the resulting GLRT decision rule depend on the signature vectors. Therefore, the complexity of this GLRT decision rule is still exponential in K. To avoid this problem we resort to the simpler least-squares (LS) approach to channel estimation. The LS channel estimate is given by

$$\hat{\mathbf{h}}_{\text{eff}} \triangleq \arg \min_{\mathbf{h}_{\text{eff}}} \{\|\mathbf{r} - \mathbf{\Phi}\mathbf{h}_{\text{eff}}\|_2^2\} = \mathbf{\Phi}^H\mathbf{r}. \quad (11)$$

Now, the GLRT decision rule can be expressed as

$$\hat{\mathbf{\Phi}} = \arg \min_{\mathbf{\Phi} \in \{\mathbf{\Phi}_0, \mathbf{\Phi}_1\}} \{\|\mathbf{r} - \mathbf{\Phi}\hat{\mathbf{h}}_{\text{eff}}\|_2^2\} = \arg \max_{\mathbf{\Phi} \in \{\mathbf{\Phi}_0, \mathbf{\Phi}_1\}} \{\|\mathbf{\Phi}^H\mathbf{r}\|_2^2\}, \quad (12)$$

where all irrelevant terms have been dropped. The FC output $u_0 = -1$ if $\hat{\mathbf{\Phi}} = \mathbf{\Phi}_0$, and $u_0 = 1$ if $\hat{\mathbf{\Phi}} = \mathbf{\Phi}_1$. Clearly, the GLRT decision rule does not require CSI and the FC does not have to know the signature vectors of the sensors.

3.3. *Performance analysis for GLRT decision rule*

For the optimum ML decision rule, a closed-form performance analysis does not seem to be feasible. However, fortunately such an analysis is possible for the more practical GLRT decision rule. In particular, the BER can be expressed as

$$P_e = P(u_0 = 1 \mid H_0)P(H_0) + P(u_0 = -1 \mid H_1)P(H_1). \quad (13)$$

Since the considered signaling scheme is symmetric in H_0 and H_1, (13) can be simplified to $P_e = P(u_0 = 1|H_0)$. Expanding now $P(u_0 = 1|H_0)$ leads to

$$P_e = \sum_{\mathcal{H}_0, \mathcal{H}_1} P(u_0 = 1 \mid \mathcal{H}_0, \mathcal{H}_1)P(\mathcal{H}_0, \mathcal{H}_1 \mid H_0), \quad (14)$$

where $P(u_0 = 1 \mid \mathcal{H}_0, \mathcal{H}_1)$ denotes the probability that $u_0 = 1$ is detected assuming that $u_k = -1$ for $k \in \mathcal{H}_0$ and $u_k = 1$ for $k \in \mathcal{H}_1$, and $P(\mathcal{H}_0, \mathcal{H}_1 \mid H_0)$ is given in Section 3.1. Exploiting the orthogonality of $\boldsymbol{\Phi}_0$ and $\boldsymbol{\Phi}_1$ and using (6) and (12), $P(u_0 = 1 \mid \mathcal{H}_0, \mathcal{H}_1)$ can be expressed as

$$P(u_0 = 1 \mid \mathcal{H}_0, \mathcal{H}_1) = P(\Delta < 0 \mid \mathcal{H}_0, \mathcal{H}_1), \qquad (15)$$

where

$$\begin{aligned} \Delta &\triangleq \|\mathbf{x}\|_2^2 - \|\mathbf{y}\|_2^2, \\ \mathbf{x} &\triangleq \sqrt{E}\mathbf{G}_{\mathcal{H}_0}\mathbf{h}_{\mathcal{H}_0} + \boldsymbol{\Phi}_0^H\mathbf{n}, \\ \mathbf{y} &\triangleq \sqrt{E}\mathbf{G}_{\mathcal{H}_1}\mathbf{h}_{\mathcal{H}_1} + \boldsymbol{\Phi}_1^H\mathbf{n}. \end{aligned} \qquad (16)$$

Since Δ is a quadratic form of Gaussian random variables, the Laplace transform $\Phi_\Delta(s)$ of the pdf of Δ can be obtained as

$$\Phi_\Delta(s) = \frac{1}{\prod_{i=1}^N (1 + s\lambda_{x_i})\prod_{i=1}^N (1 - s\lambda_{y_i})}, \qquad (17)$$

where λ_{x_i} and λ_{y_i} denote the eigenvalues of the $N \times N$ matrices

$$\begin{aligned} \mathbf{D}_x &\triangleq \varepsilon\{\mathbf{x}\mathbf{x}^H\} = E\mathbf{G}_{\mathcal{H}_0}\mathbf{G}_{\mathcal{H}_0}^H + \sigma_n^2\mathbf{I}_N, \\ \mathbf{D}_y &\triangleq \varepsilon\{\mathbf{y}\mathbf{y}^H\} = E\mathbf{G}_{\mathcal{H}_1}\mathbf{G}_{\mathcal{H}_1}^H + \sigma_n^2\mathbf{I}_N, \end{aligned} \qquad (18)$$

respectively. Thus, $P(u_0 = 1 \mid \mathcal{H}_0, \mathcal{H}_1)$ can be calculated from [19]

$$P(u_0 = 1 \mid \mathcal{H}_0, \mathcal{H}_1) = \frac{1}{2\pi j}\int_{c-j\infty}^{c+j\infty}\frac{\Phi_\Delta(s)}{s}\mathrm{d}s, \qquad (19)$$

where c is a small positive constant in the region of convergence of the integral. The integral in (19) can be either computed numerically using Gauss-Chebyshev quadrature rules [19] or exactly using [20, 21]

$$P(u_0 = 1 \mid \mathcal{H}_0, \mathcal{H}_1) = -\sum_{\text{RHS poles}}\text{Residue}\left[\frac{\Phi_\Delta(s)}{s}\right], \qquad (20)$$

where RHS stands for the right-hand side of the complex plane. The BER at the FC for the GLRT decision rule can be readily obtained by combining (14) and (19).

4. OPTIMIZATION OF CENSORING THRESHOLD d

Since a closed-form calculation of the optimum decision/censoring threshold d which minimizes P_e does not seem to be possible, we derive here a gradient algorithm for recursive optimization of d. This algorithm is given by [22]

$$d[i+1] = d[i] + \delta\frac{\partial P_e}{\partial d[i]}, \qquad (21)$$

where i is the discrete iteration index and δ is the adaptation step size. Using (14) the gradient in (21) can be expressed as

$$\frac{\partial P_e}{\partial d} = \sum_{\mathcal{H}_0, \mathcal{H}_1} P(u_0 = 1 \mid \mathcal{H}_0, \mathcal{H}_1)\frac{\partial P(\mathcal{H}_0, \mathcal{H}_1 \mid H_0)}{\partial d}, \qquad (22)$$

where we have used the fact that $P(u_0 = 1 \mid \mathcal{H}_0, \mathcal{H}_1)$ is independent of d and the remaining partial derivative is given by

$$\begin{aligned} \frac{\partial P(\mathcal{H}_0, \mathcal{H}_1 \mid H_0)}{\partial d} &= |\mathcal{S}|P_s^{|\mathcal{S}|-1}P_c^{|\mathcal{H}_0|}P_w^{|\mathcal{H}_1|}\frac{\partial P_s}{\partial d} \\ &+ |\mathcal{H}_0|P_s^{|\mathcal{S}|}P_c^{|\mathcal{H}_0|-1}P_w^{|\mathcal{H}_1|}\frac{\partial P_c}{\partial d} \\ &+ |\mathcal{H}_1|P_s^{|\mathcal{S}|}P_c^{|\mathcal{H}_0|}P_w^{|\mathcal{H}_1|-1}\frac{\partial P_w}{\partial d}. \end{aligned} \qquad (23)$$

Using (3), (4) and the fundamental theorem of calculus [23], the derivatives in (23) can be expressed as

$$\begin{aligned} \frac{\partial P_w}{\partial d} &= -\frac{1}{\sqrt{2\pi}\sigma}e^{-(d+1)^2/2\sigma^2}, \\ \frac{\partial P_c}{\partial d} &= -\frac{1}{\sqrt{2\pi}\sigma}e^{-(d-1)^2/2\sigma^2}, \\ \frac{\partial P_s}{\partial d} &= \frac{1}{\sqrt{2\pi}\sigma}\left[e^{-(d+1)^2/2\sigma^2} + e^{-(d-1)^2/2\sigma^2}\right]. \end{aligned} \qquad (24)$$

For $d = 0$, we have $|\mathcal{S}| = 0$ and since $\partial P_w/\partial d < 0$ and $\partial P_c/\partial d < 0$ we obtain $\partial P_e/\partial d < 0$. On the other hand, for $d \to \infty$, we get $|\mathcal{H}_0| \to 0$ and $|\mathcal{H}_1| \to 0$ which results in $\partial P_e/\partial d > 0$.[4] Therefore, by the mean value theorem, $\partial P_e/\partial d = 0$ is valid for at least one value of $0 \le d < \infty$ corresponding to at least one local minimum of P_e [23]. Although numerical evidence shows that there is exactly one local minimum (which therefore is also the global minimum), we cannot formally prove this due to the complexity of the involved expressions. Nevertheless, the above considerations suggest that we initialize the gradient algorithm with $d[0] = 0$ corresponding to the case of no censoring. The solution found by the algorithm is then guaranteed to yield a performance not worse than that of the no censoring case. Numerical examples will be given in the next section.

We note that d will typically be calculated at the FC and the value of d has to be conveyed to the sensors over a feedback channel. However, this feedback channel can be very low rate assuming that the statistical properties of the forward channel and the sensors vary only slowly with time.

5. SIMULATION RESULTS

In this section, we provide some numerical and simulation results for the proposed censored DSTBCs and the system model introduced in Section 2. We assume that $T = 8$ symbol intervals are available for transmission of one information bit, that is, orthogonal matrices $\boldsymbol{\Phi}_0$ and $\boldsymbol{\Phi}_1$ can be found for $N \le 4$. Here, we consider $N = 1$, $N = 2$, and $N = 4$, and generate $\boldsymbol{\Phi}_0$ and $\boldsymbol{\Phi}_1$ from the 8×8 Hadamard matrix $\mathbf{H}_8$, where the orthogonal columns of $\mathbf{H}_8$ are normalized to unit length. For example, for $N = 2\boldsymbol{\Phi}_0$ consists of the first two columns of $\mathbf{H}_8$, whereas $\boldsymbol{\Phi}_1$ consists of the third and fourth

---

[4] In fact, it can be shown that $\partial P_e/\partial d$ approaches zero from above if $d \to \infty$ corresponding to the maximum BER of $P_e = 0.5$.

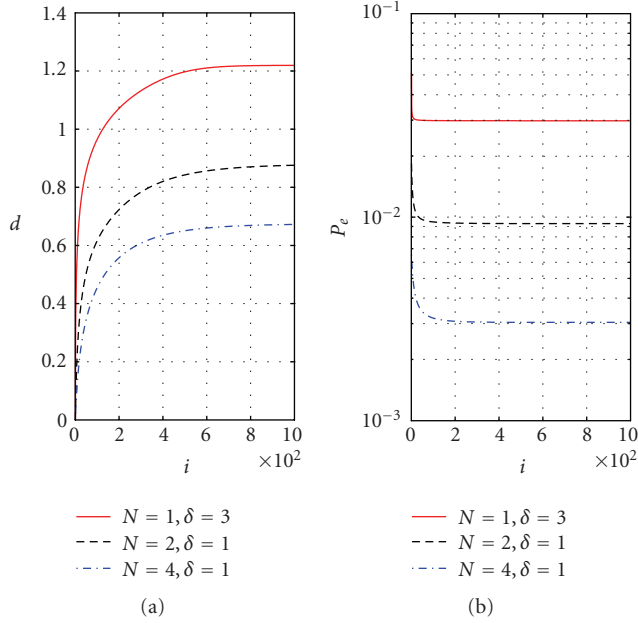FIGURE 2: d and P_e versus iteration number i for a WSN with $K = 30$ sensors using DSTBCs with $N = 1, 2,$ and 4. $10 \log_{10}(E_b/N_0) = 15$ dB, $\sigma^2 = 1/4$.

FIGURE 3: P_e versus $10 \log_{10}(E_b/N_0)$ for a WSN with $K = 30$ sensors using DSTBCs with $N = 1, 2,$ and 4. Considered cases: error-free local sensor decisions ($\sigma^2 = 0$, $d = 0$), noisy sensor decisions without censoring ($\sigma^2 = 1/4$, $d = 0$), and noisy sensor decisions with optimum censoring ($\sigma^2 = 1/4$, $d = d_{\mathrm{opt}}$).

column of $\mathbf{H}_8$. For the set of signature vectors $\mathcal{G}$, we adopted the gradient sets described in [14]. Unless stated otherwise, the sensors have a local noise variance of $\sigma^2 = 1/4$ corresponding to a signal-to-noise ratio (SNR) of 6 dB and we assume the suboptimum GLRT decision rule and P_e at the FC are obtained using the analytical results presented in Section 3.3.[5]

d and P_e versus i. First, we investigate the behavior of the adaptive algorithm described in Section 4 for optimization of d. Figure 2 shows d and the corresponding BER P_e at the FC as a function of the iteration number i for $N = 1, 2,$ and 4, respectively. The considered WSN had $K = 30$ sensors and the channel SNR was $10 \log_{10}(E_b/N_0) = 15$ dB. $d[i]$ was initialized with 0 and the step size parameter was chosen to achieve a fast convergence while avoiding instabilities. As can be observed from Figure 2 the adaptive algorithm significantly improves the BER over the iterations. While d itself requires more than 600 iterations to converge to the final optimum value, P_e does practically not change after more than 180 iterations for all considered cases. It is interesting to note that the optimum value for d decreases with increasing N, that is, for larger N less censoring should be applied. The reason for this behavior is that the maximum achievable diversity order of a DSTBC is N (cf. [14]) and therefore, the performance of the DSTBC improves notably with increasing number of transmitting sensors only until N sensors transmit. If more than N sensors transmit, the diversity order does not further improve and only a small additional coding gain can be real-
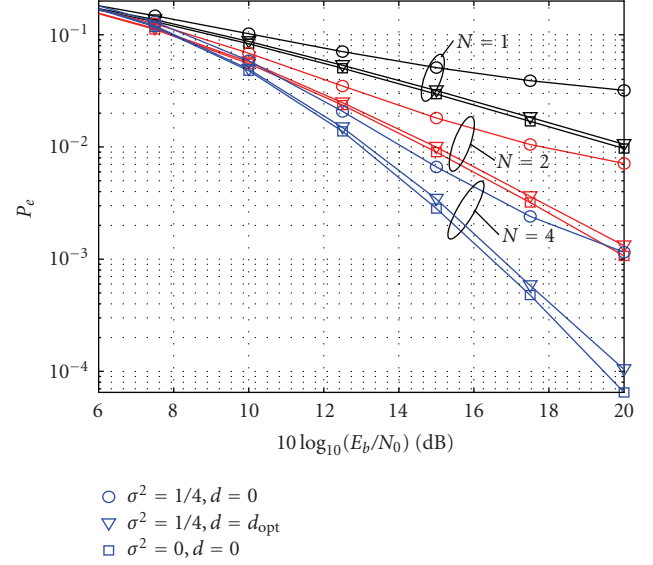
ized. On the other hand, less censoring means that more erroneous decisions are forwarded to the FC which may negate the additional coding gain.

P_e versus $10 \log_{10}(E_b/N_0)$. In Figure 3, we consider the BER achievable with the proposed censored DSTBCs at the FC of a WSN with $K = 30$ sensors as a function of the channel SNR $10 \log_{10}(E_b/N_0)$. For each considered N, we compare the BER for error-free local sensor decisions ($\sigma^2 = 0$, $d = 0$), noisy sensor decisions without censoring ($\sigma^2 = 1/4$, $d = 0$), and noisy sensor decisions with censoring ($\sigma^2 = 1/4$, $d = d_{\mathrm{opt}}$), where d_{opt} denotes the optimum decision/censoring threshold found with the gradient algorithm. Figure 3 clearly shows that DSTBCs suffer from a significant performance degradation due to erroneous decisions if censoring is not applied. Fortunately, with censoring this performance degradation can be avoided and a performance close to that of error-free local decisions can be achieved. Figure 3 also nicely illustrates the diversity gain that can be realized with censored DSTBCs.

P_e versus K. In Figure 4, we investigate the dependence of the BER on the total number of sensors in the network for $10 \log_{10}(E_b/N_0) = 15$ dB. In particular, we show in Figure 4 the BER for error-free local sensor decisions and the GLRT decision rule at the FC ($\sigma^2 = 0$, $d = 0$), noisy sensor decisions with censoring and the GLRT decision rule at the FC ($\sigma^2 = 1/4$, $d = d_{\mathrm{opt}}$), and noisy sensor decisions with censoring and the ML decision rule at the FC ($\sigma^2 = 1/4$, $d = d_{\mathrm{opt}}$).[6]

---

[5] We note that we confirmed the analytical BER results for the GLRT decision rule presented in Section 3.3 by simulations. However, we do not show the simulation results here for conciseness.

[6] We note that we use for the ML decision rule also the decision/censoring threshold d_{opt} found by the proposed gradient algorithm which is based on the GLRT decision rule. Therefore, this threshold is not strictly optimum for the ML decision rule.
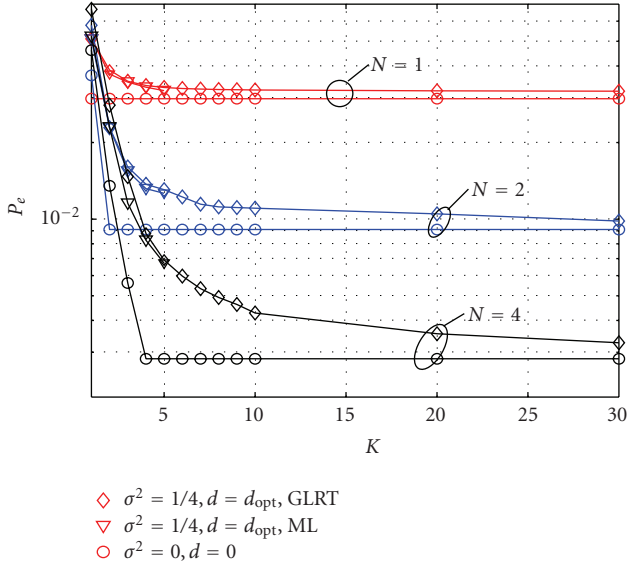
FIGURE 4: P_e versus total number of sensors K for aWSN using DST-BCs with $N = 1$, 2, and 4. $10 \log_{10}(E_b/N_0) = 15$ dB. Numerical results for error-free local sensor decisions and GLRT decision rule ($\sigma^2 = 0$, $d = 0$), numerical results for noisy sensor decisions with censoring and GLRT decision rule ($\sigma^2 = 1/4$, $d = d_{\text{opt}}$), and simulation results for noisy sensor decisions with censoring and ML decision rule ($\sigma^2 = 1/4$, $d = d_{\text{opt}}$).

FIGURE 5: P_e and d versus N for aWSN with K sensors. $\sigma^2 = 1/4$ and $10 \log_{10}(E_b/N_0) = 15$ dB. GLRT fusion rule is shown for all K (solid curves) and ML fusion rule is shown for $K = 1$ and 2 (dashed curve).

The results for the GLRT decision rule were obtained numerically based on the analytical results in Section 3.3, whereas Monte Carlo simulation was used to obtain the results for the ML decision rule. For complexity reasons, for the latter case, we only show the results for $K \leq 5$. For error-free local sensor decisions, BER is constant for $K > N$ since the diversity order is limited to N and the DSTBC achieves the same performance as the related STBC $\mathcal{C}$ for colocated antennas if all $K > N$ sensors transmit. The censored DSTBC with noisy sensor decisions approaches the performance of the DSTBC with error-free sensor decisions as the number of sensors increases. This is due to the fact that as K increases the decision/censoring threshold d_{opt} increases making the transmission of erroneous sensor decisions less likely. Figure 4 also shows that the GLRT decision rule is almost optimum and only small additional gains are possible if the significantly more complex ML decision rule is used.

P_e and d versus N. Assuming the GLRT decision rule and $10 \log_{10}(E_b/N_0) = 15$ dB at the FC, Figure 5 shows P_e and the corresponding optimum decision threshold d as a function of N for $K = 1$, 2, 4, 10 and 30. Similar to the observation we made in Figure 2, d decreases for increasing signature vector length N for all K. As we have mentioned before, the maximum achievable diversity order for DSTBC is N. For a given K, a smaller d allows more sensors to be active and thus exploits the the extra diversity benefit provided by the longer signature vectors. This figure also shows that d increases for increasing K. This can be also explained easily. For a given d and N, increasing K allows more sensors to transmit. However, our scheme only requires a certain number of sensors to be active to exploit the full diversity benefit and
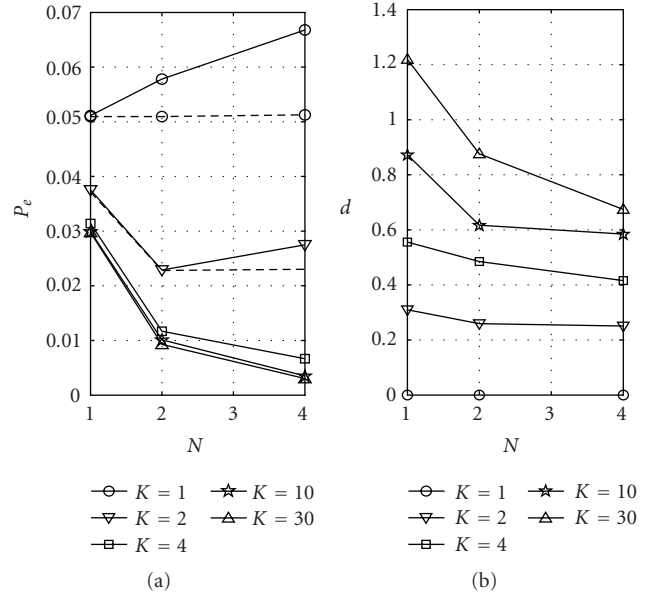
achieve a certain target BER. On the other hand, increasing d decreases the chance of having erroneous decisions being transmitted to the FC. This suggests that our scheme tries to maximize the performance by only allowing the minimum number of sensors (with quality decisions) to transmit. Finally, it is interesting to see that the P_e performance actually deteriorates for $N > K$ for the GLRT fusion rule. This is because for $N > K$ the GLRT fusion rule implicitly estimates the $N \times 1$ effective channel vector $\hat{\mathbf{h}}_{\text{eff}}$ in a noisy environment (cf. (11)) whereas the underlying channel vectors, $\mathbf{h}_{\mathcal{H}_0}$ and $\mathbf{h}_{\mathcal{H}_1}$, have a smaller dimensionality K. The increased dimensionality causes a larger channel estimation error while no diversity benefit is achieved because the maximum diversity order is limited to K [14]. In light of this degradation for the GLRT fusion rule, we also simulated the ML fusion rule for $K = 1$ and $K = 2$ (dashed curves) and clearly, as expected, the ML decision rule does not suffer from the same degradation. We note that in the practically more relevant case of $N < K$ ML and GLRT decision rules have similar performances (cf. Figure 4).

P_e and d versus SNR of local sensors. We investigate the effect of local sensor observation noise on the P_e performance in Figure 6. In particular, we plot P_e versus the SNR of local sensors $10 \log_{10}(1/\sigma^2)$ for different K and N. We assume the GLRT fusion rule at the FC and the corresponding optimum decision threshold d is also depicted. Furthermore, the channel SNR is fixed to $10 \log_{10}(E_b/N_0) = 15$ dB for all cases. As expected, the network with $K = 30$ sensors performs better than the network with $K = 10$ sensors for any N regardless of the sensor observation noise. However, this gain is minimal for large sensor SNR. This is because as the sensor SNR
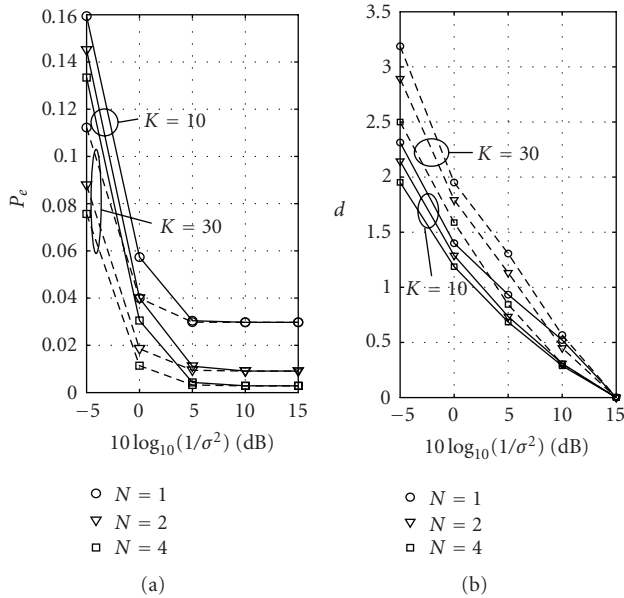
FIGURE 6: P_e and d versus $10 \log_{10}(1/\sigma^2)$ for a WSN with $K = 10$, and 30 sensors and DSTBC with $N = 1, 2$, and 4. $10 \log_{10}(E_b/N_0) = 15$ dB.
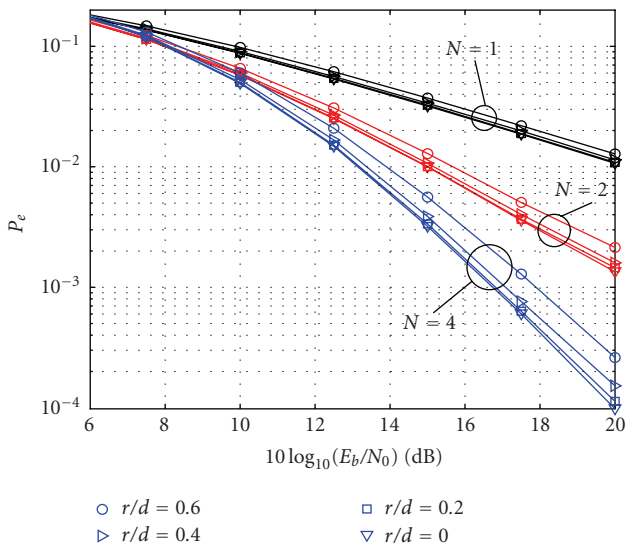

FIGURE 7: P_e versus $10 \log_{10}(E_b/N_0)$ for a WSN with $K = 30$ sensors using DSTBCs with $N = 1, 2$, and 4. $\sigma^2 = 1/4$ and i.n.d. Rayleigh fading channels.

increases, most of the sensor decisions will be correct and less censoring is required. This phenomenon is clearly supported by the corresponding d versus $10 \log_{10}(1/\sigma^2)$ figure where the optimum decision threshold d approaches zero for increasing sensor SNR. In addition, as more sensors transmit, the maximum achievable diversity order N and the channel SNR will be the ultimate factors which determine P_e and therefore, for a given N, the BER curves for $K = 10$ and $K = 30$ converge to the same value for large local sensor SNR.

I.n.d. Rayleigh fading. Until now, we have been considering i.i.d. Rayleigh fading channels. In our last example, we consider independent and nonidentically distributed (i.n.d.) fading channels. In particular, we consider a network with $K = 30$ sensors and the sensor nodes are uniformity distributed in a circle with radius r and the distance from the center of the circle to the FC is d. We assume i.n.d. Rayleigh fading between the sensors and the FC and the received power decreases as $d_k^{-\alpha}$, where d_k is the distance measured from sensor k to the FC and $\alpha = 3$ is the path loss exponent. Figure 7 depicts the simulated P_e versus $10 \log_{10}(E_b/N_0)$ for different r/d ratios. For a given N, the decision threshold d was optimized for $r/d = 0$ (corresponding to i.i.d. fading) and it was then used also for $r/d > 0$. It can be seen from the figure that, as expected, P_e increases with increasing r/d. It is also interesting to note that the performance degradation is larger for larger N. This can be explained as follows. For a given network size K, as we have seen in Figures 4 and 5, d decreases for increasing N. Since a smaller censoring threshold d corresponds to a larger number of active sensors, more sensors are negatively affected by the i.n.d. channels resulting in the greater performance degradation for larger N.

6. CONCLUSION

In this paper, we have considered the application of noncoherent DSTBCs in WSNs. We have introduced censoring as an efficient method to overcome the negative effects of erroneous local sensor decisions on the performance of the noncoherent DSTBC. Furthermore, we have derived optimum ML and suboptimum GLRT FC decision rules, and we have analyzed the performance of the latter decision rule. Based on this analysis, we have devised a gradient algorithm for recursive optimization of the decision/censoring threshold. Numerical and simulation results have shown the effectiveness of censoring which eliminates the effect of local decision errors for practically relevant BERs if the number of sensors in the network K is greater than the length of the signature vectors N or in other words, if there are enough sensors to exploit the diversity benefit provided by the DSTBC. Finally, our results have shown that the suboptimum GLRT fusion rule performs very close to the optimum ML fusion rule while having a very low complexity and allowing noncoherent detection at the FC.

REFERENCES

[1] R. R. Tenney and N. R. Sandell Jr., "Detection with distributed sensors," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 17, no. 4, pp. 501–510, 1981.

[2] J. N. Tsitsiklis, "Decentralized detection," in *Advances in Statistical Signal Processsing*, vol. 2, pp. 297–344, JAI Press, Greenwich, Conn, USA, 1993.

[3] R. Viswanathan and P. K. Varshney, "Distributed detection with multiple sensors—part I: fundamentals," *Proceedings of the IEEE*, vol. 85, no. 1, pp. 54–63, 1997.

[4] R. S. Blum, S. A. Kassam, and H. V. Poor, "Distributed detection with multiple sensors—part II: advanced topics," *Proceedings of the IEEE*, vol. 85, no. 1, pp. 64–79, 1997.

[5] P. K. Varshney, *Distributed Detection with Multiple Sensors*, Springer, Berlin, Germany, 1997.

[6] J.-J. Xiao and Z.-Q. Luo, "Universal decentralized detection in a bandwidth-constrained sensor network," *IEEE Transactions on Signal Processing*, vol. 53, no. 8, pp. 2617–2624, 2005.

[7] B. Chen, R. Jiang, T. Kasetkasem, and P. K. Varshney, "Channel aware decision fusion in wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 52, no. 12, pp. 3454–3458, 2004.

[8] R. Niu, B. Chen, and P. K. Varshney, "Fusion of decisions transmitted over Rayleigh fading channels in wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 54, no. 3, pp. 1018–1027, 2006.

[9] R. Jiang and B. Chen, "Fusion of censored decisions in wireless sensor networks," *IEEE Transactions on Wireless Communications*, vol. 4, no. 6, pp. 2668–2673, 2005.

[10] C. Rago, P. Willett, and Y. Bar-Shalom, "Censoring sensors: a low-communication-rate scheme for distributed detection," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 32, no. 2, pp. 554–568, 1996.

[11] T. Ohtsuki, "Performance analysis of statistical STBC cooperative diversity using binary sensors with observation noise," in *Proceedings of the 62nd IEEE Vehicular Technology Conference (VTC '05)*, vol. 3, pp. 2030–2033, Dallas, Tex, USA, September 2005.

[12] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1451–1458, 1998.

[13] J. N. Laneman and G. W. Wornell, "Distributed space-time-coded protocols for exploiting cooperative diversity in wireless networks," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2415–2425, 2003.

[14] S. Yiu, R. Schober, and L. Lampe, "Distributed space-time block coding," *IEEE Transactions on Communications*, vol. 54, no. 7, pp. 1195–1206, 2006.

[15] B. S. Mergen and A. Scaglione, "Randomized space-time coding for distributed cooperative communication," in *Proceedings of IEEE International Conference on Communications (ICC '06)*, vol. 10, pp. 4501–4506, Istanbul, Turkey, June 2006.

[16] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: efficient protocols and outage behavior," *IEEE Transactions on Information Theory*, vol. 50, no. 12, pp. 3062–3080, 2004.

[17] B. M. Hochwald and T. L. Marzetta, "Unitary space-time modulation for multiple-antenna communications in Rayleigh flat fading," *IEEE Transactions on Information Theory*, vol. 46, no. 2, pp. 543–564, 2000.

[18] X. Li, M. Chen, and W. Liu, "Application of STBC-encoded cooperative transmissions in wireless sensor networks," *IEEE Signal Processing Letters*, vol. 12, no. 2, pp. 134–137, 2005.

[19] E. Biglieri, G. Caire, G. Taricco, and J. Ventura-Traveset, "Computing error probabilities over fading channels: a unified approach," *European Transactions on Telecommunications*, vol. 9, no. 1, pp. 15–25, 1998.

[20] J. K. Cavers and P. Ho, "Analysis of the error performance of trellis-coded modulations in Rayleigh-fading channels," *IEEE Transactions on Communications*, vol. 40, no. 1, pp. 74–83, 1992.

[21] E. B. Saff and A. D. Snider, *Fundamentals of Complex Analysis for Mathematics, Science and Engineering*, Prentice Hall, Upper Saddle River, NJ, USA, 2nd edition, 1993.

[22] J. Nocedal and S. J. Wright, *Numerical Optimization*, Springer, New York, NY, USA, 1999.

[23] R. A. Adams, *Single Variable Calculus*, Addison-Wesley, Reading, Mass, USA, 1995.