Combining Channel Estimation and Sensor Fault Protection in Wireless Sensor Networks

Si-Yao Huang*, Chia-Lung Wu§, Po-Ning Chen*, Tsang-Yi Wang† and Yunghsiang S. Han‡
*Dept. of Electrical Eng., National Chiao-Tung Univ., Taiwan, ROC
Email: SYHuang0114@gmail.com, poning@faculty.nctu.edu.tw
§Mobile Devices Inc., Taiwan, ROC
Email: clwu@banyan.cm.nctu.edu.tw
†Institute of Comm. Eng., National Sun Yat-sen Univ., Taiwan, ROC
Email: tcwang@mail.nsysu.edu.tw
‡Dept. of Electrical Eng., National Taiwan Univ. of Sci. and Tech., Taiwan, ROC
Email: yshan@mail.ntust.edu.tw

***Abstract*—The ongoing development of wireless sensor networks (WSNs) demands not only low-power sensors and less system cost but also good performance. Considering this background, investigating a new technology to satisfy both requirements is an important issue for current development of wireless sensor network systems.**

In this paper, we consider the situation that the sensor nodes in WSNs are deployed in a harsh environment such that both channel fading and unexpected sensor faults may occur. This work then proposes to combine channel estimation and sensor-fault protection using error-correcting coding technique so as to free the costly devices of channel estimation and equalization from fusion centers. Simulations, performed to examine the performance of our proposed blind-detection scheme, show that it can compete with the conventional training-sequence-based fusion in performance when the training sequences are retained for information bearing. By this new approach, the complexity in fusion centers can be reduced and the transmitted power of sensors decreased without sacrificing the performance.

I. Introduction

A wireless sensor network (WSN) contains a number of sensors to detect the environmental variation. The sensor decisions are then transmitted to the base station. The base station, serving as a fusion center, subsequently determines what phenomenon has occurred after the collection of local decisions from sensor nodes. In real applications, the system cost is often required to be inexpensive with acceptably good performance. How to simultaneously achieve these two goals is seemingly a challenge for the design of a WSN.

When a WSN is deployed in a harsh environment, the wireless links between sensors and the fusion center may suffer channel fading. Since resources such as energy and computation power are often limited at sensor nodes, it turns out that only simple local classification can be implemented to combat such fading effect. In particular, when some sensor nodes are cornered at the margin of the target environment to be sensed, their local decisions could become nearly irrelevant to the true phenomenon. Sensor nodes are also prone to damage owing to random deployment. Placing them in a harsh or inaccessible area may make them irreplaceable. Consequently, one of the requirements for WSNs in certain applications is the capability to function *robustly* under unexpected *sensor faults*, which are ideally defined as that the local decision sent by the faulty sensor is independent of the true phenomenon [1]–[4]. It was then shown in [5], [6] that in channels without fading, the incorporation of the error-control coding technique to the design of WSNs can provide the desired robustness against sensor faults that are unaware to the fusion center.

One technique that can perhaps be used to prevent the performance degradation from faulty senor nodes is to transmit a sequence of training bits for faulty sensor detection, after which the decisions sent from faulty nodes could be ignored at the fusion center. In an environment with channel fading, these training bits could also serve their conventional role, i.e., being the basis for channel estimation. Nevertheless, the training-sequence-based technique can capture only the abnormality of local transmission modules but not the failure of computation modules or the inconsistent sensing of a node. Hence, when the local classification of a sensor happens to be independent of the phenomenon, its local decision that sends after a sequence of correctly transmitted training bits still degrade the global decision at the fusion center. Alternative technique is therefore required to prevent the performance degradation from such faulty nodes.

Recently, some works for point-to-point communications [7]–[10] proposed to use computer-searched signals (codes or constellations) for combined channel estimation and error protection in quasi-static fading channels. After the establishment of a search criterion, their signals were then searched by using the random search technique or gradient descent method. They found under the presumption of a fixed effective code rate that the best signals result in apparently better performance than a benchmark system with a certain number of training bits.

Motivated by these works, the codes that combine channel estimation and sensor fault protection are developed in this paper. Upon the reception of coded transmissions from sensors, the fusion center performs the optimal blind detection about the true phenomenon. Simulations showed that without the effort of channel estimation and equalization, the developed

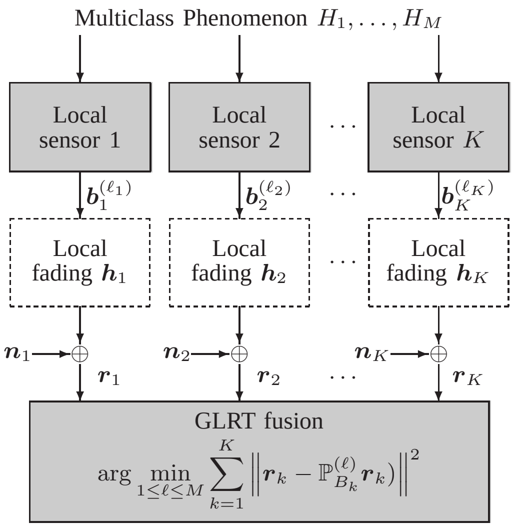

Fig. 1. System model for wireless sensor networks using error correcting codes that combine channel estimation and sensor-fault protection.

scheme can outperform the conventional training-sequence-based fusion at most SNR ranges in both fault-free and faulty WSNs. The advantage of this technique is evident: since channel estimation and equalization are implicitly implied by coding technique, no channel estimation and equalization devices are necessary in system implementation and hence the system cost can be reduced.

The remainder of the paper is organized as follows. The system model is given in Section II. The code search criterion for the proposed blind-detection scheme is presented in Section III. We investigate the performance of the proposed blind-detection scheme and compare it with the conventional training-sequence-based detection scheme in Section IV. We conclude this paper in Section V.

II. System Model

Consider a WSN with M environmental hypotheses, H_1, $H_2, \cdots, H_M$, and K sensor nodes as shown in Fig. 1. A sensor node, according to a pre-designed codebook, will decide the most probable hypothesis H_ℓ based on its local observation and transmit the codeword of length N corresponding to the favored hypothesis H_ℓ.

Let $\boldsymbol{b}_i^{(\ell)} \triangleq [b_{i,1}^{(\ell)} \quad b_{i,2}^{(\ell)} \quad \cdots \quad b_{i,N}^{(\ell)}]^T$ represent the transmitted codeword of sensor i, favoring hypothesis H_ℓ, where $b_{i,j}^{(\ell)} \in \{\pm 1\}$. Denote by $\boldsymbol{h}_i = [h_{i,1} \quad \cdots \quad h_{i,P}]^T$ the channel coefficients experienced by sensor i, and assume $\boldsymbol{h}_i$ constant during the transmission of $\boldsymbol{b}_i^{(\ell)}$. The received vector $\boldsymbol{r}_i$ due to the codeword transmission from sensor i is then given by

$$\begin{aligned} \boldsymbol{r}_i &= \begin{bmatrix} h_{i,1} & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & & \ddots & \vdots \\ h_{i,P} & \cdots & h_{i,1} & 0 & \cdots & 0 \\ 0 & h_{i,P} & \cdots & h_{i,1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & & \ddots & 0 \\ 0 & \cdots & 0 & h_{i,P} & \cdots & h_{i,1} \\ \vdots & \ddots & & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & h_{i,P} \end{bmatrix}_{L\times N} \boldsymbol{b}_i^{(\ell)} + \boldsymbol{n}_i \\ &= \mathbb{H}_i \boldsymbol{b}_i^{(\ell)} + \boldsymbol{n}_i, \end{aligned} \tag{1}$$

where $L \triangleq N+P-1$, and $\{\boldsymbol{n}_i\}_{i=1}^K$ that are assumed independent and identically distributed both spatially and temporally are the zero-mean complex Gaussian noises. Equation (1) can be straightforwardly re-written as

$$\boldsymbol{r}_i = \mathbb{B}_i^{(\ell)} \boldsymbol{h}_i + \boldsymbol{n}_i, \tag{2}$$

where

$$\mathbb{B}_i^{(\ell)} = \begin{bmatrix} b_{i,1}^{(\ell)} & 0 & \cdots & 0 \\ b_{i,2}^{(\ell)} & b_{i,1}^{(\ell)} & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ b_{i,P}^{(\ell)} & b_{i,P-1}^{(\ell)} & \cdots & b_{i,1}^{(\ell)} \\ \vdots & \ddots & \ddots & \vdots \\ b_{i,N}^{(\ell)} & b_{i,N-1}^{(\ell)} & & b_{N-P+1}^{(\ell)} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & b_{i,N}^{(\ell)} \end{bmatrix}_{L\times P} . \tag{3}$$

Based upon the received vectors, the joint maximum-likelihood (JML) decision for the true hypothesis equals

$$(\hat{\ell}, \hat{\boldsymbol{h}}_k) = \arg \min_{1\le\ell\le M} \min_{\boldsymbol{h}_k \in \mathbb{C}^P} \sum_{k=1}^{K} \left\| \boldsymbol{r}_k - \mathbb{B}_k^{(\ell)} \boldsymbol{h}_k \right\|^2,$$

where $\mathbb{C}$ denotes the set of all complex numbers. After replacing $\boldsymbol{h}_k$ with its best estimation for given ℓ, i.e.,

$$\hat{\boldsymbol{h}}_k = [(\mathbb{B}_k^{(\ell)})^T \mathbb{B}_k^{(\ell)}]^{-1} (\mathbb{B}_k^{(\ell)})^T \boldsymbol{r}_k,$$

we obtain the well-known general likelihood-ratio test (GLRT) decoding metric

$$\hat{\ell} = \arg \min_{1\le\ell\le M} \sum_{k=1}^{K} \left\| \boldsymbol{r}_k - \mathbb{P}_{B_k}^{(\ell)} \boldsymbol{r}_k \right\|^2,$$

where $\mathbb{P}_{B_k}^{(\ell)} = \mathbb{B}_k^{(\ell)} [(\mathbb{B}_k^{(\ell)})^T \mathbb{B}_k^{(\ell)}]^{-1} (\mathbb{B}_k^{(\ell)})^T$.

The decoding complexity of the GLRT usually grows exponentially with respect to the codeword length because, in general, one needs to use an exhaustive search among all candidate $\mathbb{P}_{B_k}^{(\ell)}$. Yet, when the number of hypotheses (or equivalently, the number of codewords) M is small, exhaust search fusion is still a practical choice for the fusion center design, and hence, is adopted in our WSN system.

III. Code Search Criterion under Rayleigh Fadings

In this section, we derive the criterion to be minimized during the code search process.

The detection error P_e at the fusion center can be generally given by

$$P_e = \sum_{m=1}^{M} \Pr(H_m) \Pr(\hat{\ell} \neq m | H_m).$$

Then, by denoting the local hypothesis classification at node i by s_i, we obtain

$$\begin{aligned}
&\Pr(\hat{\ell} \neq m | H_m) \\
&= \sum_{\ell_1=1}^{M} \sum_{\ell_2=1}^{M} \cdots \sum_{\ell_K=1}^{M} \Pr(s_1 = H_{\ell_1}, \cdots, s_K = H_{\ell_K} | H_m) \\
&\quad \times \Pr(\hat{\ell} \neq m | H_m, s_1 = H_{\ell_1}, \cdots, s_K = H_{\ell_K}) \\
&= \sum_{\ell_1=1}^{M} \sum_{\ell_2=1}^{M} \cdots \sum_{\ell_K=1}^{M} P^{(1)}_{\ell_1|m} P^{(2)}_{\ell_2|m} \cdots P^{(K)}_{\ell_K|m} \\
&\quad \times \Pr(\hat{\ell} \neq m | H_m, s_1 = H_{\ell_1}, \cdots, s_K = H_{\ell_K}),
\end{aligned} \tag{4}$$

where $P^{(i)}_{\ell|m}$ is the probability of node i favoring hypothesis H_ℓ when H_m is the true hypothesis. The probability $\Pr(\hat{\ell} \neq m | H_m, s_1 = H_{\ell_1}, \cdots, s_K = H_{\ell_K})$ can be further derived as

$$\begin{aligned}
&\Pr(\hat{\ell} \neq m | H_m, s_1 = H_{\ell_1}, \cdots, s_K = H_{\ell_K}) \\
&= \Pr(\hat{\ell} \neq m | s_1 = H_{\ell_1}, \cdots, s_K = H_{\ell_K}) \qquad (5) \\
&\leq \Pr\left(\sum_{k=1}^{K} \left\| \boldsymbol{r}_k - \mathbb{P}^{(m)}_{B_k} \boldsymbol{r}_k \right\|^2 \geq \min_{1 \leq \ell \leq M} \sum_{k=1}^{K} \left\| \boldsymbol{r}_k - \mathbb{P}^{(\ell)}_{B_k} \boldsymbol{r}_k \right\|^2 \right. \\
&\qquad \left. \Bigg| s_1 = H_{\ell_1}, \cdots, s_K = H_{\ell_K} \right) \\
&\leq \sum_{\ell=1, \ell \neq m}^{M} \Pr\left(\sum_{k=1}^{K} \left(\left\| \boldsymbol{r}_k - \mathbb{P}^{(m)}_{B_k} \boldsymbol{r}_k \right\|^2 - \left\| \boldsymbol{r}_k - \mathbb{P}^{(\ell)}_{B_k} \boldsymbol{r}_k \right\|^2 \right) \right. \\
&\qquad \left. \geq 0 \Bigg| s_1 = H_{\ell_1}, \cdots, s_K = H_{\ell_K} \right) \\
&\leq \sum_{\ell=1, \ell \neq m}^{M} \sum_{k=1}^{K} \Pr\left(\boldsymbol{r}_k^H \left(\mathbb{P}^{(\ell)}_{B_k} - \mathbb{P}^{(m)}_{B_k} \right) \boldsymbol{r}_k \geq 0 \right| \\
&\qquad s_1 = H_{\ell_1}, \cdots, s_K = H_{\ell_K}), \qquad (6)
\end{aligned}$$

where (5) follows from the Markovian relation among the following three items, i.e., (i) the true hypothesis H_m, (ii) the transmitted codewords at local sensors and (iii) the received vectors at the fusion center. Observe that under $s_k = H_{\ell_k}$, the covariance matrix of $\boldsymbol{r}_k$ is equal to $\mathbb{S}_{\boldsymbol{r}_k} = \mathbb{B}^{(\ell_k)}_k \mathbb{S}_{\boldsymbol{h}_k} (\mathbb{B}^{(\ell_k)}_k)^T + \sigma^2_{n_k} \mathbb{I}_L$, where $\mathbb{S}_{\boldsymbol{h}_k}$ is the covariance matrix of $\boldsymbol{h}_k$, and $\mathbb{I}_L$ is the $L \times L$ identity matrix. Thus, by denoting the eigenvalues and eigenvectors of the real and symmetric matrix $\mathbb{S}^{1/2}_{\boldsymbol{r}_k} (\mathbb{P}^{(\ell)}_{B_k} - \mathbb{P}^{(m)}_{B_k}) \mathbb{S}^{1/2}_{\boldsymbol{r}_k}$ by $\{\lambda_{n,k} = \lambda_{n,k}(\ell, m, \ell_k)\}_{n=1}^{L}$ and $\{\boldsymbol{q}_{n,k} = \boldsymbol{q}_{n,k}(\ell, m, \ell_k)\}_{n=1}^{L}$ respectively, we have

$$\mathbb{S}^{1/2}_{\boldsymbol{r}_k} (\mathbb{P}^{(\ell)}_{B_k} - \mathbb{P}^{(m)}_{B_k}) \mathbb{S}^{1/2}_{\boldsymbol{r}_k} = \sum_{n=1}^{L} \lambda_{n,k} \boldsymbol{q}_{n,k} \boldsymbol{q}^T_{n,k}.$$

Here, the dependence on ℓ, m and ℓ_k in these notations are omitted for convenience. It immediately follows that

$$\begin{aligned}
&\boldsymbol{r}_k^H (\mathbb{P}^{(\ell)}_{B_k} - \mathbb{P}^{(m)}_{B_k}) \boldsymbol{r}_k \\
&= \left(\mathbb{S}^{-1/2}_{\boldsymbol{r}_k} \boldsymbol{r}_k \right)^H \mathbb{S}^{1/2}_{\boldsymbol{r}_k} (\mathbb{P}^{(\ell)}_{B_k} - \mathbb{P}^{(m)}_{B_k}) \mathbb{S}^{1/2}_{\boldsymbol{r}_k} \left(\mathbb{S}^{-1/2}_{\boldsymbol{r}_k} \boldsymbol{r}_k \right) \\
&= \sum_{n=1}^{L} \lambda_{n,k} \left| X_{n,k} \right|^2,
\end{aligned} \tag{7}$$

where $X_{n,k} = \boldsymbol{q}^T_{n,k} \mathbb{S}^{-1/2}_{\boldsymbol{r}_k} \boldsymbol{r}_k$. Note that under the premise that $\{\boldsymbol{h}_k\}_{k=1}^{K}$ and $\{\boldsymbol{n}_k\}_{k=1}^{K}$ are independent zero-mean complex Gaussian random vectors with independent components, and that they are independent to each other, $X_{n,k}$ becomes zero-mean complex Gaussian distributed with unit variance and is independent for different n. This implies that $\{|X_{n,k}|^2\}_{1 \leq n \leq L}$ are independent χ^2-distributed with two degree of freedom. We then assume without loss of generality that

1) there are $\bar{L}_k$ different non-zero eigenvalues in $\{\lambda_{n,k}\}_{1 \leq n \leq L}$, denoted by $\{\bar{\lambda}_{\ell,k}\}_{\ell=1}^{\bar{L}}$ with $\bar{\lambda}_{1,k} > \bar{\lambda}_{2,k} > \cdots > \bar{\lambda}_{\bar{L},k}$;
2) the individual orders of multiplicity of $\{\bar{\lambda}_{\ell,k}\}_{\ell=1}^{\bar{L}}$ are $\{o_{\ell,k}\}_{\ell=1}^{\bar{L}}$.

As a result, the sum of (7) can be rewritten as

$$\begin{aligned}
&\boldsymbol{r}_k^H (\mathbb{P}^{(\ell)}_{B_k} - \mathbb{P}^{(m)}_{B_k}) \boldsymbol{r}_k \\
&= \sum_{n=1}^{L} \lambda_{n,k} \left| X_{n,k} \right|^2 = \sum_{\ell=1}^{\bar{L}_k} \bar{\lambda}_{\ell,k} \cdot \chi^2(2 o_{\ell,k}),
\end{aligned}$$

where $\chi^2(\nu)$ is χ^2-distributed with ν degree of freedom. It can be further derived that

$$\begin{aligned}
&E\left[\exp\left\{ \mathtt{i} t \boldsymbol{r}_k^H (\mathbb{P}^{(\ell)}_{B_k} - \mathbb{P}^{(m)}_{B_k}) \boldsymbol{r}_k \right\} \right] \\
&= \prod_{n=1}^{\bar{L}_k} \left(1 - 2\mathtt{i} \bar{\lambda}_{n,k} t \right)^{-o_{n,k}/2},
\end{aligned} \tag{8}$$

where $\mathtt{i} = \sqrt{-1}$ denotes the imaginary unit. Resuming the derivation in (6) using (8), we obtain

$$\begin{aligned}
&\Pr(\hat{\ell} \neq m | H_m, s_1 = H_{\ell_1}, \cdots, s_K = H_{\ell_K}) \\
&\leq \sum_{\ell=1, \ell \neq m}^{M} \sum_{k=1}^{K} \frac{1}{2\pi} \int_0^{\infty} \\
&\quad \left(\int_{-\infty}^{\infty} \prod_{n=1}^{\bar{L}_k} \left(1 - 2\mathtt{i} \bar{\lambda}_{n,k} t \right)^{-o_{n,k}/2} e^{-\mathtt{i} t r} dt \right) dr.
\end{aligned} \tag{9}$$

In the special case of Rayleigh fading channels, where $\{\boldsymbol{h}_k\}_{k=1}^{K}$ are assumed zero-mean Gaussian distributed, (9) can be further derived as (see [12])

$$\Pr(\hat{\ell} \neq m | H_m, s_1 = H_{\ell_1}, \cdots, s_K = H_{\ell_K}) \leq \sum_{\ell=1,\ell\neq m}^{M} \sum_{k=1}^{K} \sum_{n=1}^{\bar{L}_k} \frac{1}{(o_{n,k}-1)!} \left[\frac{\partial^{(o_{n,k}-1)}}{\partial x^{(o_{n,k}-1)}} F_{n,k}(x) \right]_{x=\bar{\lambda}_{n,k}}$$

where

$$F_{n,k}(x) = x^{\sum_{i=1}^{\bar{L}_k} o_{i,k}-1} \prod_{u=1,u\neq n}^{\bar{L}_k} (x-\bar{\lambda}_{u,k})^{-o_{u,k}}.$$

We finally conclude

$$P_e \leq \sum_{m=1}^{M} \Pr(H_m) \sum_{\ell_1=1}^{M} \cdots \sum_{\ell_K=1}^{M} P^{(1)}_{\ell_1|m} \cdots P^{(K)}_{\ell_K|m} \sum_{\ell=1,\ell\neq m}^{M} \sum_{k=1}^{K} \sum_{n=1}^{\bar{L}_k} \frac{1}{(o_{n,k}-1)!} \left[\frac{\partial^{(o_{n,k}-1)}}{\partial x^{(o_{n,k}-1)}} F_{n,k}(x) \right]_{x=\bar{\lambda}_{n,k}} \quad (10)$$

In this work, the code design is to search the codebook $\{\boldsymbol{b}_i^{(\ell)}\}_{1\leq i\leq K, 1\leq \ell\leq M}$, where $\boldsymbol{b}_i^{(\ell)} \in \{\pm 1\}^N$, which minimizes the upper bound in (10). In order to find a feasible solution within an acceptable simulation time, the simulated annealing algorithm [11] is employed for this nonlinear and NP-hard optimization problem (which belongs to the so-called binary integer programming). It is a popular algorithm for finding the statistically near-optimum solution of certain problems. It mimics a physical phenomenon about temperature control. Specifically, a system containing high entropy at high temperature is set at the beginning. Then, the system temperature is cooled down gradually at each iteration until a stable solution at low temperature is reached. The error probability bound in (10) is now regarded as the system entropy. Thus, during the procedure, the error probability bound is gradually cooled down until the target codebook is found.

IV. Simulation Results

In this section, we examine the performance of the found code that combines channel estimation and sensor-fault protection in WSNs when $M = 2$ and $K = 5$. In our simulations, the channel coefficients $\boldsymbol{h}_i$ is taken to be independently and identically zero-mean complex Gaussian distributed with memory order $(P-1) = 1$ and $E[|h_{i,1}|^2] = E[|h_{i,2}|^2] = 1/2$, and $\{\boldsymbol{h}_i\}_{i=1}^K$ are also independent across nodes. Equally likely hypothesis is assumed. The operating signal-to-noise ratio (SNR) for the codeword search is set to be 10 dB. Since the blind detection at the fusion center cannot differentiate the transmissions between $\boldsymbol{b}$ and $-\boldsymbol{b}$, we fix b_1 as -1 in the code design.

Our scheme is compared with the conventional training-sequence-based fusion. It is known that the least square (LS) estimate of the channel coefficients for a given training sequence $\boldsymbol{t}_k$ for sensor k is given by (see [13])

$$\hat{\boldsymbol{h}}_k = (\mathbb{T}_k^T \mathbb{T}_k)^{-1} \mathbb{T}_k^T \boldsymbol{y}_k,$$

where the relation between $\boldsymbol{t}_k$ and $\mathbb{T}_k$ is exactly the same as that between $\boldsymbol{b}_k^{(\ell)}$ and $\mathbb{B}_k^{(\ell)}$ (cf. (3)), and $\boldsymbol{y}_k$ is the reception

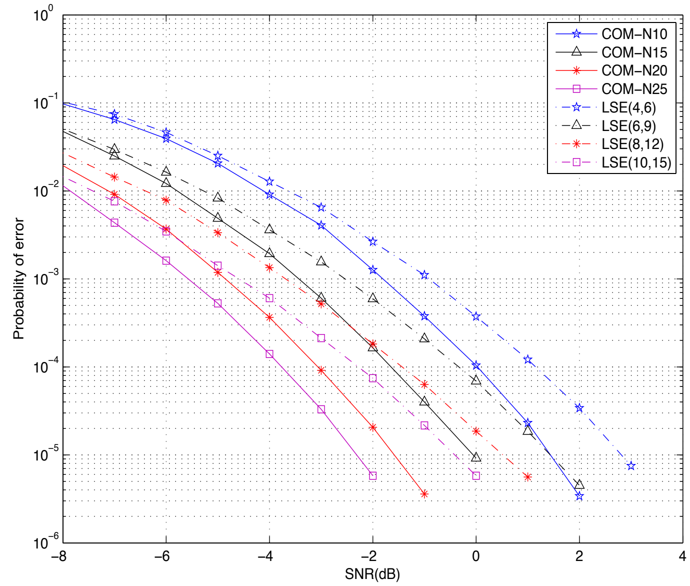

Fig. 2. Performances of hypothesis detection using soft-decision fusion. Five sensors and two hypotheses are assumed in this figure. LSE(x, y) represents that the LSE scheme uses x training bits and y code bits. COM-Nz represents that the codeword length of the COM scheme is z (bits).

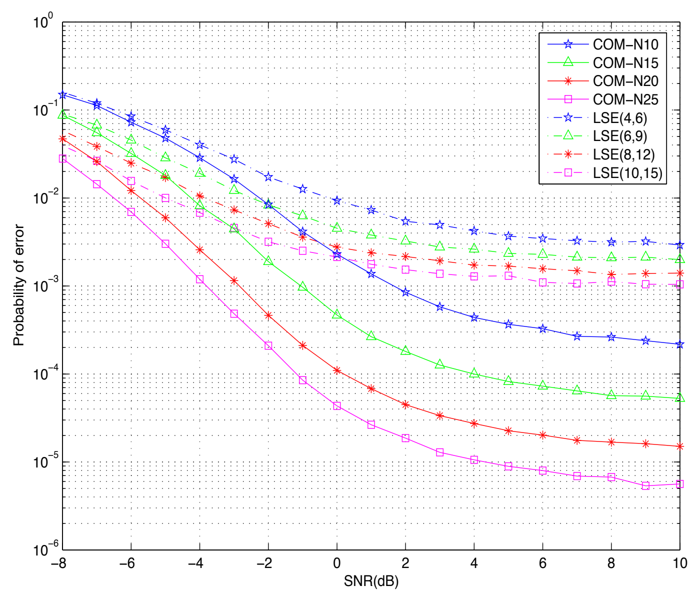

Fig. 3. Performance of hypothesis detection using soft-decision fusion. Five sensors, two hypotheses, and one faulty sensor are assumed in this figure. LSE(x, y) represents that the LSE scheme uses x training bits and y code bits. COM-Nz represents that the codeword length of the COM scheme is z (bits).

due to the transmission of $\boldsymbol{t}_k$. We can then use the estimated channel coefficients to determine the hypothesis as

$$\hat{\ell} = \arg \min_{1\leq \ell \leq M} \sum_{k=1}^{K} \left\| \boldsymbol{r}_k - \mathbb{B}_k^{(\ell)} \hat{\boldsymbol{h}}_k \right\|^2 .$$

Notably, in the above formula, $\mathbb{B}_k^{(\ell)}$ should have a smaller matrix size than $L \times P = (N + P - 1) \times P$ since we assume for fair comparison that the total number of training bits and code bits should remain N.

Through simulations, we found that taking $D = 0.4 \times N$ as the length of training sequence will yield the best performance. We will accordingly adopt this ratio when the hypothesis-detection-based-on-LS-channel-estimate scheme (abbreviated as the LSE scheme) is referred to. Since the number of hypotheses is two, it is straightforward that for the LSE scheme, the optimal code under perfectly known channel coefficients should be one consisting of all-(-1) and all-$(+1)$ codewords.

The combined-channel-estimation-and-sensor-fault-protection scheme that we proposed is abbreviated as the COM scheme. Since what we concern in this paper is the effect of combined channel estimation and sensor-fault protection in comparison with the conventional training-sequence-based fusion, we assume perfect local classification, i.e., $P_{\ell|\ell}^{(i)} = 1$, in our simulations.

Figure 2 presents the detection error probabilities of both the LSE and the COM schemes for binary hypotheses testing. The results indicate that in a fault-free situation, the COM scheme always performs better than the respective equal-length LSE scheme. In fact, the COM scheme can outperform the LSE scheme of longer length at medium to high SNRs. For example, the probability of error for the COM-N20 scheme is lower than that of the LSE$(10, 15)$ scheme when the SNR is greater than -6 dB, and the length of the latter is actually 5 bits longer than that of the former. Similar phenomena can be observed between the COM-N10 and the LSE$(6, 9)$, as well as between the COM-N15 and the LSE$(8, 12)$. We can then conclude that the COM scheme is indeed more power-efficient than the LSE scheme by transmitting less number of bits but achieving better performance at medium to high SNRs.

Next, we re-perform the simulations in Fig. 2 by introducing one faulty sensor and summarize the results in Fig. 3. After transmitting the training sequence, the faulty sensor will transmit sequence $\{\pm 1\}^{N-D}$ randomly, independent of its local observation. Our result indicates that the COM schemes again outperform the LSE schemes. In addition, the error floors for both COM and LSE schemes appear due to that the fusion center is unaware of which sensor is faulty. Because the information-bearing bits of the LSE schemes are shorter than those of the COM schemes, the error floor of the LSE schemes is considerably higher than the respective COM schemes.

V. Conclusion

In this work, the scheme of combined channel estimation and sensor-fault protection in WSNs has been examined. As the training sequence is retained for information-bearing, our simulations indicate that the proposed coding scheme can compete with the conventional training-sequence-based fusion in performance without the effort of channel estimation and equalization.

References

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam and E. Cayirci, "A survey on sensor networks," *IEEE Communications Magazine*, pp. 102–114, August 2002.

[2] D. Culler, D. Estrin and M. Srivastava, "Overview of sensor networks," *IEEE Computer*, vol. 37, no. 8, pp. 41–49, August 2004.

[3] L. Dan, K. D. Wong, H. H. Yu and A. M. Sayeed, "Detection, classification, and tracking of targets," *IEEE Trans. Signal Processing*, vol. 19, no. 3, pp. 17–29, March 2002.

[4] F. Sivrikaya and B. Yener, "Time synchronization in sensor networks: a survey," *IEEE Network*, vol. 18, no. 4, pp. 45–50, July–Auguest 2004.

[5] C. Yao, P.-N. Chen, T.-Y. Wang, Y. S. Han and P. K. Varshney, "Performance analysis and code design for minimum hamming distance fusion in wireless sensor networks," *IEEE Trans. Inform. Theory*, vol. 53, no. 5, pp. 1716–1734, May 2007.

[6] T.-Y. Wang, Y. S. Han, B. Chen and P. K. Varshney, "A combined decision fusion and channel coding scheme for distributed fault-tolerant classification in wireless sensor networks," *IEEE Trans. Wireless Comun.*, vol. 5, no. 7, pp. 1695–1705, July 2006.

[7] M. Skoglund and S. Parkvall, "Code design for combined channel estimation and error protection," *IEEE Trans. Inform. Theory*, vol. 48, no. 2, pp. 1162–1171, May 2002.

[8] O. Coskun and K. M. Chugg, "Combined coding and training for unknown ISI channels," *IEEE Trans. Commun.*, vol. 53, no. 8, pp. 1310-1322, August 2005.

[9] M. Beko, J. Xavier and V. A. N. Barroso, "Noncoherent communication in multiple-antenna system: receiver design and codebook construction," *IEEE Trans. Signal Process.*, vol. 55, no. 12, pp. 5703-5715, Dec. 2007.

[10] J. Giese and M. Skoglund, "Single- and Multiple-Antenna Constellations for Communication Over Unknown Frequency-Selective Fading Channels" *IEEE Trans. Inform. Theory*, vol. 53, no. 8, pp. 1584-1594, 2007, .

[11] S. Kirkpatrick, C. D. Gelatt Jr. and M. P. Vecchi, "Optimization by simulated annealing," *Science*, vol. 220, no. 4598, pp. 671–680, 1983.

[12] J. P. Imhof, "Computing the distribution of quadratic forms in normal variables," *Biometrica*, vol. 48, no. 3-4, pp. 419–426, 1961.

[13] Harry L. Van Trees, *Detection, Estimation, and Modulation Theory, Part I*, Wiley-Interscience, USA, 2001.