Performance Analysis and Code Design for Minimum Hamming Distance Fusion in Wireless Sensor Networks

Chien Yao, Po-Ning Chen, *Senior Member, IEEE*, Tsang-Yi Wang, *Member, IEEE*, Yunghsiang S. Han, *Member, IEEE*, and Pramod K. Varshney, *Fellow, IEEE*

***Abstract*—Distributed classification fusion using error-correcting codes (DCFECC) has recently been proposed for wireless sensor networks operating in a harsh environment. It has been shown to have a considerably better capability against unexpected sensor faults than the optimal likelihood fusion. In this paper, we analyze the performance of a DCFECC code with minimum Hamming distance fusion. No assumption on identical distribution for local observations, as well as common marginal distribution for the additive noises of the wireless links, is made. In addition, sensors are allowed to employ their own local classification rules. Upper bounds on the probability of error that are valid for any finite number of sensors are derived based on large deviations technique. A necessary and sufficient condition under which the minimum Hamming distance fusion error vanishes as the number of sensors tends to infinity is also established. With the necessary and sufficient condition and the upper error bounds, the relation between the fault-tolerance capability of a DCFECC code and its pair-wise Hamming distances is characterized, and can be used together with any code search criterion in finding the code with the desired fault-tolerance capability. Based on the above results, we further propose a code search criterion of much less complexity than the minimum Hamming distance fusion error criterion adopted earlier by the authors. This makes the code construction with acceptable fault-tolerance capability for a network with over a hundred of sensors practical. Simulation results show that the code determined based on the new criterion of much less complexity performs almost identically to the best code that minimizes the minimum Hamming distance fusion error. Also simulated and discussed are the performance trends of the codes searched based on the new simpler criterion with respect to the network size and the number of hypotheses.**

***Index Terms*—Coding, detection, classification, information fusion, fault tolerance, wireless sensor networks.**

Manuscript received November 5, 2005; revised January 17, 2007. This work was supported in part by the Research Center of Wireless Network and Multimedia Communication under "Aim for Top University Plan" project of NSYSU and the Ministry of Education, Taiwan, and the NSC of Taiwan, R.O.C., under Grants NSC 95-2221-E-009-247 and NSC 95-2221-E-009-563. The material in this paper was presented in part at the IEEE International Symposium on Information Theory, Seattle, WA, July 2006.

C. Yao is with Department of Electronics Engineering, National Chiao-Tung University, Hsinchu, Taiwan 300, R.O.C. (e-mail: zyao.ee88g@nctu.edu.tw).

P.-N. Chen is with the Department of Communication Engineering, National Chiao-Tung University, Hsinchu, Taiwan 300, R.O.C. (e-mail: qponing@mail.nctu.edu.tw).

T.-Y. Wang is with the Institute of Communications Engineering, National Sun Yat-Sen University, Kaohsiung, Taiwan 804, R.O.C. (e-mail: tcwang@mail.nsysu.edu.tw).

Y. S. Han is with the Graduate Institute of Communication Engineering, National Taipei University, Taipei, Taiwan 237, R.O.C. (e-mail:yshan@mail.ntpu.edu.tw).

P. K. Varshney is with the Department of Electrical Engineering and Computer Science, Syracuse University, Syracuse, NY 13244 USA (e-mail: varshney@syr.edu).

Communicated by A. Høst-Madsen, Associate Editor for Detection and Estimation.

Digital Object Identifier 10.1109/TIT.2007.894670

I. Introduction

RECENT advances in processor, radio, and memory technology have generated a great interest in the notion of deploying a large number of networked sensors for applications such as environment monitoring. The classification of target objects, as well as their tracking, are the fundamental requirements in these applications [1], [2], [8], [10], [11], [17]. In this paper, we consider a wireless sensor network (WSN) that consists of N geographically dispersed sensors, wireless (and hence noisy) one-way communication links, and a fusion center. Limitation on the communication bandwidth in wireless links due to the consideration of economical energy consumption at local sensors prevents the system from conveying raw observation data to the fusion center. A local compression on the raw observation data thus has to be employed at each sensor. Usually, the information content of the compressed outputs from local sensors is of fewer bits than $\log_2(M)$ in a WSN, where M is the number of object classes to be distinguished. In this work, we are specifically concerned with the case where the sensor nodes only send out binary decisions to the fusion center at which they are fused to produce the final M-ary decision.

Another issue that may be encountered in a WSN is that sensors are prone to be blocked or even damaged when they are deployed in a harsh environment [1]. In addition, a low-cost sensor that is manufactured by a simple technology may suffer from hardware, as well as software, malfunctions after deployment. As a result, the fault-tolerance capability to protect against unexpected sensor failures is also of equal importance to the performance and complexity of a WSN.

To fulfill the above mentioned requirements, a distributed classification fusion approach using error correcting codes (DCFECC) has been proposed to provide good fault-tolerance capability under feasible system complexity [18]. In the proposed approach, an $M \times N$ error-correcting code matrix is first designed by either simulated annealing or cyclic column replacement, where each row of N bits forms a codeword that corresponds to one of M hypotheses. Each local sensor then outputs the respective code bit of the codeword corresponding to the declared hypothesis that is locally determined based

0018-9448/$25.00 © 2007 IEEE

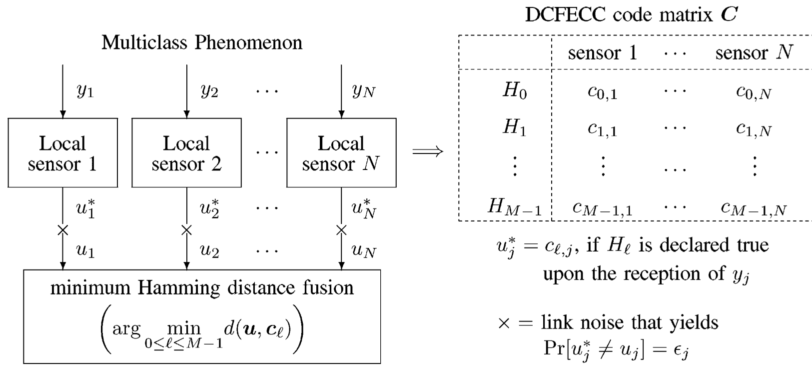

Fig. 1. System model for DCFECC.

on its own observation. Upon receipt of the binary outputs from sensors, the fusion center makes a multiclass decision based on minimum Hamming distance decoding. Unlike the conventional approach that employs the optimal maximum *a posteriori* probability (MAP) fusion rule, it was shown in [18] that with adequately high probability, the decision made by the minimum Hamming distance fusion rule can fall into the correct acceptance region even if several sensor faults are present.

In [7], we have provided the performance analysis of the minimum Hamming distance fusion approach when the number of sensors is sufficiently large. This asymptotic performance analysis for the distributed binary detection/classification problem considered here is different from those investigated by [4], [6], [16], where the MAP fusion rule rather than the minimum Hamming distance fusion rule was used.

In this work, we extend our analysis in [7] by relaxing the assumptions of common distribution for all local observations and identical local classification rule for all sensors. Also, the additive noises over the wireless links is allowed to be independent but nonidentical in statistics. Contrary to the requirement of sufficiently large number of sensors in [7], the probability bounds obtained in this work are now valid for any finite number of sensors. In particular, a necessary and sufficient condition under which the minimum Hamming distance fusion error vanishes as the number of sensors tends to infinity is established. With the necessary and sufficient condition and the upper bounds on the error probability, the relation between the fault-tolerance capability of a DCFECC code and its pair-wise Hamming distances can be analytically characterized. It can thereby be used together with any code search criterion for finding the code matrix with the desired fault-tolerance capability. Most importantly, a code search criterion of much less complexity than the minimum Hamming distance fusion error criterion adopted in [18] is proposed. Based on this, code construction with acceptable fault-tolerance capability for a network with over a hundred of sensors becomes possible. Simulation results show that the code determined based on the new criterion with much less complexity performs almost identically to the best code that minimizes the minimum Hamming distance fusion error. Also simulated and discussed are the performance trends of the codes searched based on the new simpler criterion with respect to the number of sensors and the number of hypotheses. Detailed discussions are provided in subsequent sections.

This paper is organized in the following fashion. The system model is described in detail in Section II. The error bounds are derived using large deviations technique in Section III, followed by the establishment of the necessary and sufficient condition under which the fusion error vanishes as the number of sensors N becomes large. In Section IV, we characterize the fault-tolerance capability of a DCFECC code with minimum Hamming distance fusion. With the availability of the upper bounds on the error probability and the characterization of the fault-tolerance capability, the new code search criterion is presented in Section V. Section VI summarizes and discusses the simulation and numerical results obtained in this work. Section VII concludes the paper.

For better readability, the proofs of the supportive lemmas are deferred till the Appendix.

II. System Model

As depicted in Fig. 1, we consider the distributed M-ary classification problem in a parallel fusion network, which is perhaps the topology that has received the most attention in the area of WSNs [3], [4], [9], [10], [11], [12], [14], [19]. In this problem, all the local sensors observe the same phenomenon that statistically belongs to one of the M possible classes. Independent interferences are assumed present at the local sensors, which, in mathematics, makes the local observations $\{y_j\}_{j=1}^N$ conditionally independent across sensors given each hypothesis. Also assume that each local sensor classifies its own observation, independent of all others, to one of the M hypotheses using its own decision rule. In other words, the local sensor nodes need not employ identical decision rules. We then denote by $h_{\ell|i}^{(j)}$ the probability of classifying H_ℓ given that H_i is the true hypothesis for sensor j.

After the observation is locally classified at sensor j, a local output bit u_j^* is transmitted through a noisy channel to the fusion center at which place the N received bits are combined to yield the fusion decision. Due to channel transmission errors, the word $\boldsymbol{u} = (u_1, u_2, \ldots, u_N)$ received at the fusion center may not equal the transmitted word $\boldsymbol{u}^* = (u_1^*, u_2^*, \ldots, u_N^*)$. In this paper, we assume that the event of link error, i.e., $[u_j \ne u_j^*]$, is independent for all the communication links between sensors and the fusion center, and is also independent of the observations $\{y_j\}_{j=1}^N$ as well as the true hypothesis H_i, and its probability $\Pr[u_j \ne u_j^*]$ is denoted by ϵ_j.

Ideally, the objective of the distributed classification in parallel fusion networks is to determine the local classification rules and the fusion rule such that the minimum fusion error can be achieved. For this reason, the optimal MAP fusion rule is often employed at the fusion center. However, such an optimal design was shown to degrade drastically in performance when unexpected sensor failures are present [18]. This results in the motivation behind the DCFECC codes in [18], i.e., to borrow the technique of error-correcting (namely, fault-correcting) codes to design a wireless sensor network system that is much less sensitive, and hence, more robust, to sensor faults that are unseen by the fusion center.

By following this motivation, a code matrix $\boldsymbol{C}$ of which the design methodology will be covered in Section V is specified in advance in the DCFECC system as shown on the right side of Fig. 1. This code matrix is an $M \times N$ matrix with element $c_{\ell,j} \in \{0,1\}$, where $\ell = 0,\ldots,M-1$ and $j = 1,\ldots,N$. Each hypothesis H_ℓ is associated with a row in the code matrix. Each column of $\boldsymbol{C}$ stands for the local binary outputs corresponding to the locally classified hypotheses at the respective sensor. Thus, sensor j transmits $c_{\ell,j}$, if H_ℓ is declared to be true locally. Clearly, at least $\lceil \log_2(M) \rceil$ sensor nodes are required for the identification of M hypotheses, and it requires more sensor nodes to provide the coding redundancy for fault tolerance. For notational convenience, $\boldsymbol{c}_\ell \triangleq (c_{\ell,1}, c_{\ell,2}, \ldots, c_{\ell,N})$ is used to denote the row of $\boldsymbol{C}$ corresponding to the hypothesis H_ℓ.

The redundancy that constitutes the desired fault tolerance comes from the adopted minimum distance fusion rule, or specifically, $\omega = \arg\min_{0\le\ell\le M-1} d(\boldsymbol{u},\boldsymbol{c}_\ell)$, where $d(\cdot,\cdot)$ is the Hamming distance [18]. The tie-break rule is to randomly pick a codeword from those with the same smallest Hamming distance to the received vector $\boldsymbol{u}$.

III. Performance Analysis

In this section, we first derive a large deviation probability bound for finite sample size. Based on the probability bound, we analyze the performance of the distributed M-ary classification system using minimum Hamming distance fusion. A necessary and sufficient condition under which the error rate of the DCFECC codes vanishes as the number of sensors tends to infinity is then established.

A. Large Deviation Probability Bound for Finite Sample Size

From the minimum Hamming distance fusion, i.e.,

$$\omega = \arg\min_{0\le\ell\le M-1} d(\boldsymbol{u},\boldsymbol{c}_\ell) = \arg\min_{0\le\ell\le M-1} \sum_{j=1}^{N} z_{\ell,j}$$

where "$\oplus$" denotes the exclusive–OR operation and $z_{\ell,j} \triangleq 2(u_j \oplus c_{\ell,j}) - 1$, it can be anticipated that the analysis of the system performance relies completely on the probabilities of events $[z_{\ell,j} = 1]$ and $[z_{\ell,j} = -1]$. Note that given the true hypothesis is H_i, the more negative the quantity $\sum_{j=1}^{N} z_{i,j}$, the smaller the fusion error is. This induces the necessity of finding a good probability bound for the sum of independent random variables $\{z_{i,j}\}_{j=1}^{N}$ given H_i is the true hypothesis in the following. For notational simplicity, we drop the redundant subscript i in $z_{i,j}$ in the derivation that follows.

Lemma 1: Let $\{Z_j\}_{j=1}^{\infty}$ be independent antipodal random variables with

$$\Pr[Z_j = 1] = q_j \quad \text{and} \quad \Pr[Z_j = -1] = 1 - q_j.$$

Define

$$\varphi_m(\theta) \triangleq \frac{1}{m} \log E[\exp\{\theta(Z_1 + \cdots + Z_m)\}],$$

and

$$I_m(x) \triangleq \sup_{\theta\in\Re} [\theta x - \varphi_m(\theta)].$$

Then, if $\lambda_m \triangleq E[Z_1 + \cdots + Z_m]/m < 0$

$$\begin{aligned}\Pr\{Z_1 + \cdots + Z_m \ge 0\} &\le \exp\{-m \cdot I_m(0)\} \\ &= \inf_{\theta\in\Re} \exp\left\{\sum_{j=1}^{m} \log\left(q_j e^{\theta} + (1-q_j)e^{-\theta}\right)\right\}.\end{aligned}$$

Remark: Since $\lambda_m < 0$, $I_m(0)$ remains the same if we redefine $I_m(x)$ as $\sup_{\theta\ge0}[\theta x - \varphi_m(\theta)]$. Hence, with the assumption of $\lambda_m < 0$, the result of Lemma 1 can be re-expressed as

$$\begin{aligned}&\Pr\{Z_1 + \cdots + Z_m \ge 0\} \\ &\quad \le \inf_{\theta\ge0} \exp\left\{\sum_{j=1}^{m} \log\left(q_j e^{\theta} + (1-q_j)e^{-\theta}\right)\right\}. \qquad (1)\end{aligned}$$

The probability bound in (1) does not exhibit any apparent relation with λ_m, namely, the average of the means of $\{Z_i\}_{i=1}^{m}$. This can be amended by the next lemma.

Lemma 2: If $\lambda_m \triangleq E[Z_1 + \cdots + Z_m]/m < 0$, then

$$\Pr\{Z_1 + \cdots + Z_m \ge 0\} \le (1 - \lambda_m^2)^{m/2}.$$

B. Performance Analysis for Distributed M-Ary Classification Fusion System With Minimum Hamming Distance Fusion

Based on the probability bounds obtained in the previous subsection, we can upper-bound the error probability of the distributed M-ary classification system using minimum Hamming distance fusion rule by the following theorem.

Theorem 1: Let P_e be the average probability of minimum Hamming distance fusion error given as

$$P_e \triangleq \frac{1}{M} \sum_{i=0}^{M-1} \Pr(\text{fusion decision} \ne H_i \,|\, H_i).$$

If for every $\ell \ne i$

$$\sum_{\{j\in[1,\ldots,N]: c_{\ell,j} \ne c_{i,j}\}} E[z_{i,j}] = \sum_{j=1}^{N} (c_{\ell,j} \oplus c_{i,j})(2q_{i,j} - 1) < 0 \qquad (2)$$

where $0 \leq \ell, i \leq M-1$, $z_{i,j} \triangleq 2(u_j \oplus c_{i,j}) - 1$, and $q_{i,j} \triangleq \Pr\{z_{i,j} = 1 \,|\, H_i\}$, then

$$P_e \leq \frac{1}{M} \sum_{i=0}^{M-1} \sum_{0\leq \ell \leq M-1, \ell \neq i} \inf_{\theta \geq 0} \exp\left\{ \sum_{j=1}^{N} \log\left(q_{i,j}e^{\theta} + (1-q_{i,j})e^{-\theta}\right)^{c_{\ell,j}\oplus c_{i,j}} \right\}. \quad (3)$$

Proof:

$$\begin{aligned}
&\Pr(\text{fusion decision} \neq H_i | H_i) \\
&\quad \leq \Pr\left(d(\boldsymbol{u}, \boldsymbol{c}_i) \geq \min_{0\leq \ell \leq M-1, \ell\neq i} d(\boldsymbol{u}, \boldsymbol{c}_\ell) | H_i \right) \\
&\quad \leq \sum_{0\leq \ell\leq M-1, \ell\neq i} \Pr\left(d(\boldsymbol{u}, \boldsymbol{c}_i) \geq d(\boldsymbol{u}, \boldsymbol{c}_\ell) | H_i \right) \\
&\quad = \sum_{0\leq \ell\leq M-1, \ell\neq i} \Pr\left(\sum_{\{j\in[1,\ldots,N]: c_{\ell,j}\neq c_{i,j}\}} z_{i,j} \geq 0 | H_i \right).
\end{aligned}$$

Observe that

$$\begin{aligned}
\Pr(z_{i,j} = 1 | H_i) &= \Pr(u_j \oplus c_{i,j} = 1 | H_i) \\
&= \Pr(u_j = u_j^* \quad \text{and } u_j^* \oplus c_{i,j} = 1 | H_i) \\
&\quad + \Pr(u_j \neq u_j^* \quad \text{and } u_j^* \oplus c_{i,j} = 0 | H_i) \\
&= \Pr(u_j = u_j^*) \Pr(u_j^* \oplus c_{i,j} = 1 | H_i) \\
&\quad + \Pr(u_j \neq u_j^*) \Pr(u_j^* \oplus c_{i,j} = 0 | H_i) \\
&= \epsilon_j + (1 - 2\epsilon_j) \sum_{k=0}^{M-1} (c_{i,j} \oplus c_{k,j}) h_{k|i}^{(j)} \quad (4)
\end{aligned}$$

and $\{z_{i,j}\}_{j=1}^{N}$ are independent across sensors given H_i is true. Therefore, by Lemma 1

$$\begin{aligned}
&\Pr\left(\sum_{\{j\in[1,\ldots,N]: c_{\ell,j}\neq c_{i,j}\}} z_{i,j} \geq 0 \middle| H_i \right) \\
&\qquad \leq \inf_{\theta\geq 0} \exp\left\{ \sum_{j=1}^{N} \log\left(q_{i,j}e^{\theta} + (1-q_{i,j})e^{-\theta}\right)^{c_{\ell,j}\oplus c_{i,j}} \right\}
\end{aligned}$$

which results in

$$\begin{aligned}
P_e &= \frac{1}{M} \sum_{i=0}^{M-1} \Pr(\text{decision} \neq H_i | H_i) \\
&\leq \frac{1}{M} \sum_{i=0}^{M-1} \sum_{0\leq \ell\leq M-1, \ell\neq i} \\
&\quad \inf_{\theta\geq 0} \exp\left\{ \sum_{j=1}^{N} \log\left(q_{i,j}e^{\theta} + (1-q_{i,j})e^{-\theta}\right)^{c_{\ell,j}\oplus c_{i,j}} \right\}.
\end{aligned}$$

Theorem 1 provides an upper bound on the probability of error by means of Lemma 1. Based on Lemma 2, the next corollary shows that the upper bound in (3) can be further upper-bounded by quantities that are only functions of the negative quantity defined in (2). As a result, the intuition that a DCFECC code with larger pair-wise Hamming distances is expected to perform better can be justified.

Corollary 1: Under condition (2), the average probability of minimum Hamming distance fusion error can also be bounded above by

$$\begin{aligned}
P_e &\leq \frac{1}{M} \sum_{i=0}^{M-1} \sum_{0\leq \ell\leq M-1, \ell\neq i} \left(1 \right. \quad (5) \\
&\qquad \left. - \left(\frac{\sum_{j=1}^{N} (c_{\ell,j} \oplus c_{i,j})(2q_{i,j} - 1)}{d(\boldsymbol{c}_\ell, \boldsymbol{c}_i)} \right)^2 \right)^{d(\boldsymbol{c}_\ell, \boldsymbol{c}_i)/2} \\
&\leq (M-1)\left(1 - \lambda_{\max}^2\right)^{d_{\min}/2} \quad (6)
\end{aligned}$$

where $d_{\min} \triangleq \min_{0\leq \ell, i\leq M-1, \ell\neq i} d(\boldsymbol{c}_\ell, \boldsymbol{c}_i)$, and

$$\lambda_{\max} \triangleq \max_{0\leq \ell, i\leq M-1, \ell\neq i} \frac{1}{d(\boldsymbol{c}_\ell, \boldsymbol{c}_i)} \sum_{j=1}^{N} (c_{\ell,j} \oplus c_{i,j})(2q_{i,j} - 1). \quad (7)$$

Notably, $\lambda_{\max} < 0$ by condition (2), and

$$\lambda_{\max} \geq \min_{0\leq i\leq M-1, 1\leq i\leq N} (2q_{i,j} - 1) \geq -1.$$

Proof: The proof of inequality (5) follows a similar procedure as in Theorem 1 except that Lemma 2 is used instead of Lemma 1. Inequality (6) is a direct consequence of inequality (5). □

With the above corollary, we figure that if for some $\delta > 0$, $\lambda_{\max} < -\delta$ for all sufficiently large N, the DCFECC decoding error vanishes exponentially as $d_{\min}$ approaches infinity. Since under a fixed number of hypotheses, $d_{\min}$ usually grows linearly with the number of sensors N for typical DCFECC codes, we conclude that the average error probability for the distributed M-ary classification system using minimum Hamming distance fusion can be made zero asymptotically as N goes to infinity, and the error exponent is bounded below by

$$\liminf_{N\to\infty} -\frac{1}{N} \log P_e \geq \liminf_{N\to\infty} -\frac{d_{\min}}{2N} \log(1 - \lambda_{\max}^2)$$

as long as $\limsup_{N\to\infty} \lambda_{\max} < 0$. Next, we will show that the assumption that $\limsup_{N\to\infty} \lambda_{\max} > 0$ leads to a nonvanishing P_e, and hence, establish the necessary and sufficient condition under which P_e vanishes.

Theorem 2: If $\limsup_{N\to\infty} \lambda_{\max} > 0$, P_e is bounded away from zero infinitely often in number of sensors.

Proof: The assumption that $\limsup_{N\to\infty} \lambda_{\max} > 0$ implies the existence of $\delta > 0$ such that $\lambda_{\max} > \delta$ for infinitely many N. Hence, for any N satisfying $\lambda_{\max} > \delta$, there exist $\ell = \ell(N)$ and $i = i(N)$ such that

$$\sum_{j=1}^{N} (c_{\ell,j} \oplus c_{i,j})(2q_{i,j} - 1) > \delta \cdot d(\boldsymbol{c}_\ell, \boldsymbol{c}_i). \quad (8)$$

By defining $z_{i,j}$ the same as in Theorem 1, we obtain

$$\begin{aligned}
\mu_{\ell,i} &\triangleq E\left[\sum_{\{j\in[1,\ldots,N]: c_{\ell,j}\neq c_{i,j}\}} z_{i,j} \right] \\
&= \sum_{\{j\in[1,\ldots,N]: c_{\ell,j}\neq c_{i,j}\}} E[z_{i,j}] \\
&= \sum_{j=1}^{N} (c_{\ell,j} \oplus c_{i,j})(2q_{i,j} - 1) > \delta \cdot d(\boldsymbol{c}_\ell, \boldsymbol{c}_i).
\end{aligned}$$

As a result

$$\begin{aligned}
&\Pr(\text{fusion decision} \neq H_i | H_i) \\
&\geq \Pr\left(d(\boldsymbol{u}, \boldsymbol{c}_i) > \min_{0 \leq \ell \leq M-1, \ell \neq i} d(\boldsymbol{u}, \boldsymbol{c}_\ell) \middle| H_i \right) \\
&\geq \Pr\left(d(\boldsymbol{u}, \boldsymbol{c}_i) > d(\boldsymbol{u}, \boldsymbol{c}_\ell) | H_i \right) \\
&= \Pr\left(\sum_{\{j \in [1,\ldots,N]: c_{\ell,j} \neq c_{i,j}\}} z_{i,j} > 0 \middle| H_i \right) \\
&\geq \Pr\left(\sum_{\{j \in [1,\ldots,N]: c_{\ell,j} \neq c_{i,j}\}} z_{i,j} - \mu_{\ell,i} > 0 \middle| H_i \right) \\
&= \Pr\left(\frac{\sum_{\{j \in [1,\ldots,N]: c_{\ell,j} \neq c_{i,j}\}} (z_{i,j} - E[z_{i,j}])}{\sqrt{\sum_{\{j \in [1,\ldots,N]: c_{\ell,j} \neq c_{i,j}\}} \mathrm{Var}[z_{i,j}]}} > 0 \middle| H_i \right) \\
&\to \frac{1}{2}, \text{ if } d(\boldsymbol{c}_\ell, \boldsymbol{c}_i) \text{ approaches infinity}
\end{aligned}$$

where the last step follows from the central limit theorem for sum of independent bounded variables. Thus, the claim of the theorem holds for the case that $d(\boldsymbol{c}_\ell, \boldsymbol{c}_i)$ tends to infinity.

In situations when $d(\boldsymbol{c}_\ell, \boldsymbol{c}_i)$ is bounded as N approaches infinity (in which case a bad DCFECC code design results), the theorem is trivially valid. □

The final lemma in this section shows that the upper bounds in Corollary 1, as well as the expression of $\lambda_{\max}$, can be greatly simplified for identical sensor systems.

Lemma 3: Suppose that $\epsilon_j = \epsilon$ for $1 \leq j \leq N$, where $0 \leq \epsilon < 1/2$, and $h_{k|i}^{(j)} = h_{k|i}$ is the same for all sensors. Then, if $\lambda_{\max} < 0$, we have

$$\begin{aligned}
P_e &\leq \frac{1}{M} \sum_{i=0}^{M-1} \sum_{0 \leq \ell \leq M-1, \ell \neq i} \left(1 - (1-2\epsilon)^2 \right. \\
&\quad \left. \times \left(\frac{\sum_{k=0}^{M-1} h_{k|i} [d(\boldsymbol{c}_i, \boldsymbol{c}_k) - d(\boldsymbol{c}_\ell, \boldsymbol{c}_k)]}{d(\boldsymbol{c}_\ell, \boldsymbol{c}_i)} \right)^2 \right)^{d(\boldsymbol{c}_\ell, \boldsymbol{c}_i)/2} \quad (9) \\
&\leq (M-1) \left(1 - \lambda_{\max}^2 \right)^{d_{\min}/2} \quad (10)
\end{aligned}$$

where $\lambda_{\max}$ is simplified to

$$\lambda_{\max} = (1-2\epsilon) \max_{0 \leq \ell, i \leq M-1, \ell \neq i} \frac{1}{d(\boldsymbol{c}_\ell, \boldsymbol{c}_i)} \times \sum_{k=0}^{M-1} h_{k|i} [d(\boldsymbol{c}_i, \boldsymbol{c}_k) - d(\boldsymbol{c}_\ell, \boldsymbol{c}_k)].$$

Lemma 3 indicates that under an identical sensor system assumption, the two upper bounds in Corollary 1 are reduced to functions of pair-wise Hamming distances. This greatly simplifies their evaluation.

By denoting

$$\beta_{\max} \triangleq \frac{\lambda_{\max}}{(1-2\epsilon)} = \max_{0 \leq \ell, i \leq M-1, \ell \neq i} \frac{1}{d(\boldsymbol{c}_\ell, \boldsymbol{c}_i)} \times \sum_{k=0}^{M-1} h_{k|i} [d(\boldsymbol{c}_i, \boldsymbol{c}_k) - d(\boldsymbol{c}_\ell, \boldsymbol{c}_k)]$$

we observe that $\lambda_{\max}$ is given by the product of $(1-2\epsilon)$ and $\beta_{\max}$, where the former only depends on the noises corresponding to communication links between sensors and the fusion center, while the latter is only a function of the local classification accuracy and the adopted DCFECC code. The upper bound in (10) can then be rewritten as $(M-1)(1-(1-2\epsilon)^2\beta_{\max}^2)^{d_{\min}/2}$. Hence, the effects of link noise and local classification accuracy can be separately considered through the help of bound (10).

For antipodal transmission over additive white Gaussian noise (AWGN) channels, we have $\epsilon = \frac{1}{2}\mathrm{erfc}(\sqrt{\gamma_s})$, where $\mathrm{erfc}(\cdot)$ is the complementary error function, and γ_s is the signal-to-noise ratio of the communication link. Fig. 2 then shows that the error bound in (10) reaches its ultimate floor value $7(1-\beta_{\max}^2)^{d_{\min}/2}$ when γ_s is larger than 7 dB, which corresponds to $\epsilon = 0.0007727$, and this threshold is independent of the local classification accuracy and the DCFECC codes adopted.

We can similarly characterize the effect of local classification accuracy through the help of the simple probability bound in (10). A usual assumption on the statistics of local observation y_j is that y_j is Gaussian distributed with mean ℓ and variance $1/\gamma_o$ given that hypothesis H_ℓ is true. Define the local classification rule as H_i is declared true if

$$(y_j - i)^2 \leq \min_{0 \leq \ell \leq M-1, \ell \neq i} (y_j - \ell)^2.$$

Then

$$h_{k|i} = \begin{cases} \rho_{k-i}, & \text{if } k = 0 \\ \rho_{k-i} - \rho_{k-i-1}, & \text{if } 1 \leq k \leq M-2 \\ 1 - \rho_{k-i-1}, & \text{if } k = M-1 \end{cases}$$

where $\rho_k = \Phi((k+0.5)\sqrt{\gamma_o})$, and $\Phi(\cdot)$ is the standard normal cumulative distribution function. As a result, $h_{k|i}$ approaches 1 as $\gamma_o \uparrow \infty$ for $k = i$, and 0, otherwise, and

$$\lim_{\gamma_o \uparrow \infty} \beta_{\max} = \lim_{\gamma_o \uparrow \infty} \max_{0 \leq \ell, i \leq M-1, \ell \neq i} \frac{1}{d(\boldsymbol{c}_\ell, \boldsymbol{c}_i)} \times \sum_{k=0}^{M-1} h_{k|i} [d(\boldsymbol{c}_i, \boldsymbol{c}_k) - d(\boldsymbol{c}_\ell, \boldsymbol{c}_k)] = -1.$$

Hence, the ultimate floor values of bound (10) for the codes in Table I (see Section VI-C) tend to

$$7(1-(1-2\epsilon)^2)^{d_{\min}/2} = 7(4\epsilon(1-\epsilon))^{d_{\min}/2}$$

when γ_o grows beyond a constant threshold 17 dB as depicted in Fig. 3.

IV. Fault-Tolerance Capability

The wireless sensor network considered is likely to contain faulty sensor nodes due to harsh environmental conditions. Faults may include all misbehaviors, ranging from simple *random sensor faults* or *stuck-at faults*[1] to sensors that behave

[1]By "random sensor fault," we mean that the sensor sends out 1 or 0 randomly regardless of the local observation. Also, a sensor with stuck-at-one (respectively, stuck-at-zero) fault will always transmit one (respectively, zero) to the fusion center, and neglect the local observation it sensed.

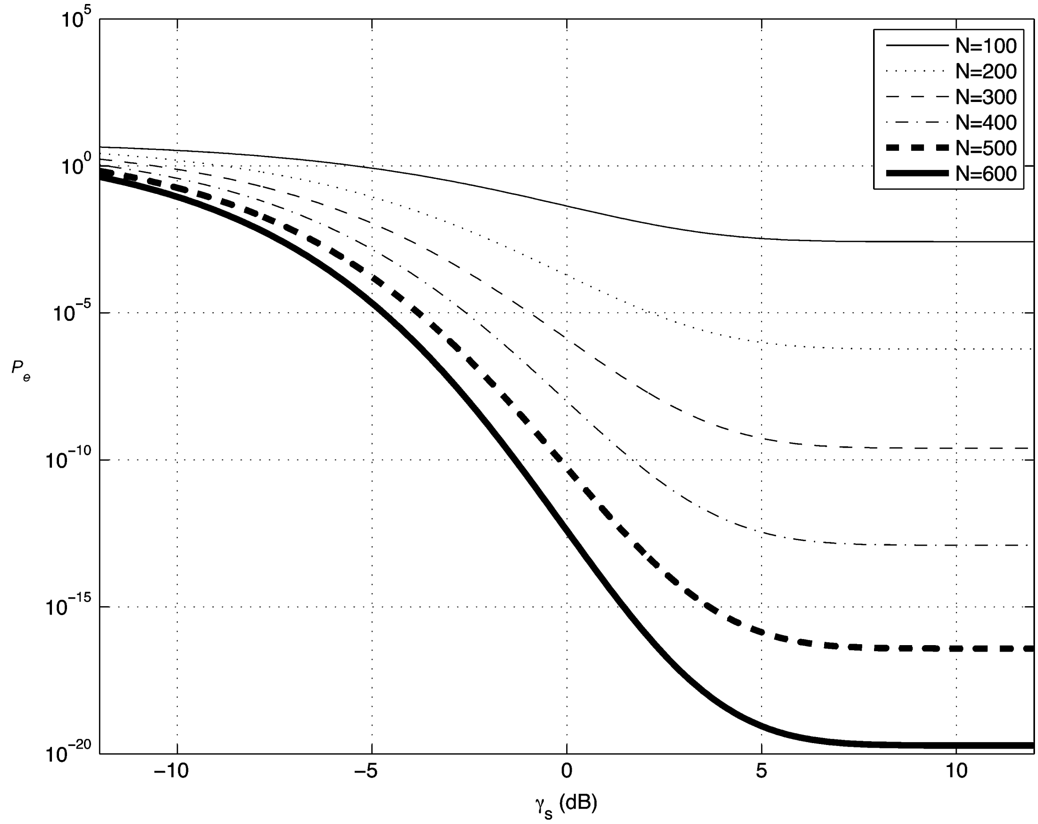

Fig. 2. Bound (10) for the codes listed in Table I (see Section VI-C) at $\gamma_o = 6$ dB, where we assume that the local observation y_j is Gaussian distributed with mean ℓ and variance $1/\gamma_o$ given hypothesis H_ℓ is true.

arbitrarily or maliciously. In this section, we remark on the fault-tolerance capability for the system employing minimum Hamming distance fusion rule according to Corollary 1.

From the upper bound in (6) and the definition of $\lambda_{\max}$ in (7), we observe that the smaller the $q_{i,j}$, the more negative $\lambda_{\max}$ is, which in turn gives a smaller upper probability bound. When sensor faults (SF) occur, $q_{i,j}$ is no longer given by (4), but becomes a function of the new statistics of u_j^* owing to sensor faults. For example, when stuck-at-one fault occurs at sensor j, $\Pr\{u_j^* = 1|H_i\} = 1$ for $0 \le i \le M-1$. Hence

$$\begin{aligned} q_{i,j}^{(\mathrm{SF})} &= \Pr(u_j = u_j^*)\Pr(u_j^* \oplus c_{i,j} = 1|H_i) \\ &\quad + \Pr(u_j \neq u_j^*)\Pr(u_j^* \oplus c_{i,j} = 0|H_i) \\ &= (1-\epsilon_j)(1-c_{i,j}) + \epsilon_j c_{i,j} \end{aligned}$$

and has nothing to do with the local classification accuracy. Similarly, for stuck-at-zero fault

$$\begin{aligned} q_{i,j}^{(\mathrm{SF})} &= \Pr(u_j = u_j^*)\Pr(u_j^* \oplus c_{i,j} = 1|H_i) \\ &\quad + \Pr(u_j \neq u_j^*)\Pr(u_j^* \oplus c_{i,j} = 0|H_i) \\ &= (1-\epsilon_j)c_{i,j} + \epsilon_j(1-c_{i,j}). \end{aligned}$$

In case a random fault occurs, in which

$$\Pr\{u_j^* = 0|H_i\} = \Pr\{u_j^* = 1|H_i\}$$

$q_{i,j}^{(\mathrm{SF})} = 1/2$. Accordingly, $q_{i,j}^{(\mathrm{SF})}$ may range from $\min\{\epsilon_j, 1-\epsilon_j\}$ to $\max\{\epsilon_j, 1-\epsilon_j\}$. As no prior information on the sensor faults type as well as the indices of faulty sensors is assumed known at the fusion center, it is reasonable to consider the fault-tolerance capability of the system by the worst case scenario in which $q_{i,j}^{(\mathrm{SF})} = \max\{\epsilon_j, 1-\epsilon_j\}$. Thus

$$\begin{aligned} q_{i,j} &= \epsilon_j + (1-2\epsilon_j)\sum_{k=0}^{M-1}(c_{i,j} \oplus c_{k,j})h_{k|i}^{(j)} \\ &\le \epsilon_j + (1-2\epsilon_j) \le \max\{\epsilon_j, 1-\epsilon_j\} = q_{i,j}^{(\mathrm{SF})} \end{aligned}$$

and sensor fault under the worst case surely degrades the system performance bound in (6). Then, the next corollary, which is a straightforward extension of Corollary 1 based on the above discussion, can be used to characterize the fault-tolerance capability of a DCFECC coding system.

Corollary 2: Let $\mathcal{F}$ be the set of indices of faulty sensors. Then, if $\lambda_{\max}(\mathcal{F}) < 0$, we have the expression at the bottom of the following page, where the superscript "c" denotes the set complement operation and

$$\begin{aligned} \lambda_{\max}(\mathcal{F}) &= \max_{0\le \ell, i \le M-1, \ell \neq i} \frac{1}{d(\boldsymbol{c}_\ell, \boldsymbol{c}_i)} \\ &\quad \times \left(\sum_{j\in\mathcal{F}^c} (c_{\ell,j} \oplus c_{i,j})(2q_{i,j}-1) \right. \\ &\quad \left. + \sum_{j\in\mathcal{F}} (c_{\ell,j} \oplus c_{i,j}) \left(2q_{i,j}^{(\mathrm{SF})} - 1\right) \right). \end{aligned}$$

By assuming that

$$q_{i,j}^{(\mathrm{SF})} = \max\{\epsilon_j, 1-\epsilon_j\}$$

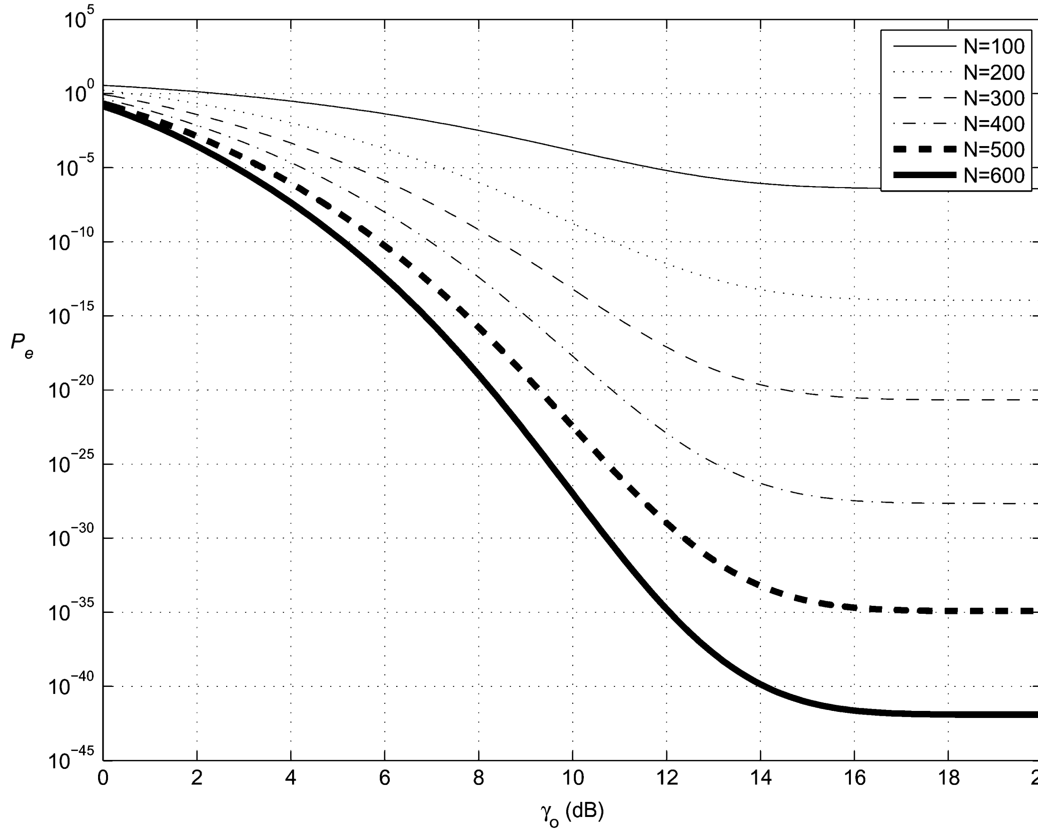

Fig. 3. Bound (10) for the codes in Table I (see Section VI-C) at $\gamma_s = 0$ dB.

we can verify based on the above corollary that

$$\begin{aligned}\lambda_{\max}(\mathcal{F}) - \lambda_{\max} &\le 2 \max_{0\le \ell, i \le M-1, \ell\neq i} \frac{1}{d(\boldsymbol{c}_\ell, \boldsymbol{c}_i)} \\ &\quad\times \sum_{j\in\mathcal{F}} (c_{\ell,j}\oplus c_{i,j})\left(q_{i,j}^{(\mathrm{SF})} - q_{i,j}\right) \\ &\le 2 \max_{0\le \ell, i \le M-1, \ell\neq i} \frac{1}{d(\boldsymbol{c}_\ell, \boldsymbol{c}_i)} \sum_{j\in\mathcal{F}} |1-2\epsilon_j| \\ &\le \frac{2}{d_{\min}} \sum_{j=1}^{|\mathcal{F}|} |1-2\epsilon_{(j)}| \qquad (11)\end{aligned}$$

where $\{\epsilon_{(j)}\}_{j=1}^{N}$ is the sorted counterpart of $\{\epsilon_j\}_{j=1}^{(N)}$, satisfying

$$\left|1-2\epsilon_{(1)}\right| \ge \left|1-2\epsilon_{(2)}\right| \ge \cdots \ge \left|1-2\epsilon_{(N)}\right|.$$

In order to guarantee that P_e vanishes, it suffices to have

$$\lambda_{\max} + \frac{2}{d_{\min}} \sum_{j=1}^{|\mathcal{F}|} |1-2\epsilon_{(j)}| < 0. \qquad (12)$$

For an identical sensor system where $\epsilon_j = \epsilon$ for $1 \le j \le N$, this condition reduces to

$$d_{\min} > -2|1-2\epsilon| \frac{|\mathcal{F}|}{\lambda_{\max}} = 2\frac{|\mathcal{F}|}{|\beta_{\max}|}. \qquad (13)$$

As

$$\begin{aligned}\lambda_{\max} &\ge \min_{0\le i\le M-1, 1\le j\le N} (2q_{i,j}-1) \\ &= -(1-2\epsilon)\left(1 - 2\sum_{k=0}^{M-1} (c_{i,j}\oplus c_{k,j}) h_{k|i}^{(j)}\right) \\ &\ge -|1-2\epsilon|\end{aligned}$$

for an identical sensor system, we have

$$d_{\min} > -2|1-2\epsilon| \frac{|\mathcal{F}|}{\lambda_{\max}} = 2\frac{|\mathcal{F}|}{|\beta_{\max}|} \ge 2|\mathcal{F}|. \qquad (14)$$

It is worth mentioning that the condition $d_{\min} > 2|\mathcal{F}|$ that was used as a code search requirement in [18] resembles the interpretation for conventional coding techniques, which states that a

$$\begin{aligned}P_e &\le \frac{1}{M}\sum_{i=0}^{M-1} \sum_{\substack{0\le \ell \le M-1\\ \ell\neq i}} \left(1 - \left(\frac{\sum_{j\in\mathcal{F}^c}(c_{\ell,j}\oplus c_{i,j})(2q_{i,j}-1) + \sum_{j\in\mathcal{F}}(c_{\ell,j}\oplus c_{i,j})(2q_{i,j}^{(\mathrm{SF})}-1)}{d(\boldsymbol{c}_\ell, \boldsymbol{c}_i)}\right)^2\right)^{\frac{d(\boldsymbol{c}_\ell,\boldsymbol{c}_i)}{2}} \\ &\le (M-1)(1-\lambda_{\max}^2(\mathcal{F}))^{d_{\min}/2}\end{aligned}$$

code with minimum pair-wise Hamming distance $d_{\min}$ can tolerate around $d_{\min}/2$ errors. However, inequality (14) hints that a larger $d_{\min}$ than $(2|\mathcal{F}|)/|\beta_{\max}|$ instead of $2|\mathcal{F}|$ may be necessary for an (identical) fault-tolerant sensor system due to the local classification inaccuracy. Thus, in the worst case, where the fusion center has no information on both the sensor fault types and indices of faulty sensors, the number of faulty sensors allowable for the codes in Table I (see Section VI-C) is only two-thirds of $d_{\min}/2$ as $\beta_{\max} = \lambda_{\max}/(1-2\epsilon) \approx -0.67$ on an average in this table. Inequality (14) also interestingly indicates that for an identical sensor system, the worst case fault-tolerance requirement has nothing to do with the link noise. Inequality (14) will reduce to the heuristic constraint of $d_{\min} > 2|\mathcal{F}|$ when all the misclassification probabilities become zero (in which case $h_{i|i} = 1$ for $0 \le i \le M-1$, and hence $\beta_{\max} = -1$ regardless of the codes adopted).

Summarizing the above discussion, we may define the *fault-tolerance capability* of a DCFECC code as the maximal number of faulty sensors allowable subject to the validity of (12) as parallel to the usual definition of error-correcting capability (i.e., $d_{\min}/2$) of block codes ([13, p. 65]). As a consequence of the definition, a sequence of DCFECC codes guarantee to have vanishing fusion error as N goes to infinity if the number of faulty sensors is always restricted within the fault-tolerance capabilities of these codes. The guarantee to have asymptotic zero-fusion-error is analogous to that in digital communication, when the number of errors is less than the error-correcting capabilities of concerned block codes, error-free transmission of these codes is guaranteed. Based on this definition, the fault-tolerance capability of a DCFECC code is equal to $\lfloor |\beta_{\max}| \times (d_{\min}/2) \rfloor$ for identical sensor systems.

In the next section, we will set the target *fault-tolerance capability* as an auxiliary constraint, in addition to the minimization of (5), during the search of a fault-tolerant DCFECC code.

V. Computer Search of DCFECC Codes

Computer search for a DCFECC code based on the minimum probability-of-fusion-error criterion for a large sensor network system is infeasible due to the prohibitive algorithmic complexity even by the simulated annealing or the cyclic column replacement algorithms recommended in [18]. However, a large sensor network that consists of either a large number of sensors or a large number of hypotheses under classification may still be encountered in practice. This raises the research issue on how to find a DCFECC code that performs well for a large sensor network.

At a first glance, one may think that the minimization of $\lambda_{\max}$ defined in Corollary 1 is a good alternative criterion for code search for a large sensor network. Yet, a simple example with $N = M = 2$, $\epsilon_1 = \epsilon_2 = \epsilon$, $h^{(1)}_{0|1} = h^{(1)}_{1|0} = 1/4$, and $h^{(2)}_{0|1} = h^{(2)}_{1|0} = 3/8$ indicates that the code matrix that minimizes the resultant

$$\lambda_{\max} = -[(1-2\epsilon)/4][2(c_{0,1} \oplus c_{1,1}) + (c_{0,2} \oplus c_{1,2})]/[(c_{0,1} \oplus c_{1,1}) + (c_{0,2} \oplus c_{1,2})]$$

should satisfy $(c_{0,1} \oplus c_{1,1}) = 1$ and $(c_{0,2} \oplus c_{1,2}) = 0$, and can be shown straightforwardly not to be the most fault-tolerant design of $(c_{0,1} \oplus c_{1,1}) = (c_{0,2} \oplus c_{1,2}) = 1$ for $M = 2$. It can be conjectured from the error bound in (6) that a good criterion to be minimized should be a function of both $\lambda_{\max}$ and $d_{\min}$. The above example shows the validity of the conjecture, and the code matrix with $(c_{0,1} \oplus c_{1,1}) = (c_{0,2} \oplus c_{1,2}) = 1$ does minimize the error bound $(1-\lambda^2_{\max})^{d_{\min}/2}$.

A better criterion than $(1-\lambda^2_{\max})^{d_{\min}/2}$, especially for a non-identical sensor network in which some sensors or some hypotheses have much larger $q_{i,j}$ than the others, is the upper probability bound in (5). A simple calculation shows that the number of exclusive-OR, multiplication and addition operations required to determine (5) for a selected code matrix $\boldsymbol{C}$ is of the order $O(NM^2)$.[2] However, the required operations for the determination of $\{\mathcal{D}_i\}_{i=0}^{M-1}$ that are necessary for the determination of the minimum Hamming distance fusion error, which can be closely approximated as

$$\frac{1}{M}\sum_{i=0}^{M-1} \Pr\{\boldsymbol{u} \in \mathcal{D}_i^c | H_i\} = \frac{1}{M}\sum_{i=0}^{M-1}\sum_{\boldsymbol{u}\in\mathcal{D}_i^c}\prod_{j=1}^{N}[u_j\bar{q}_{i,j} + (1-u_j)(1-\bar{q}_{i,j})] \quad (15)$$

where

$$\bar{q}_{i,j} = \Pr\{u_j = 1|H_i\} = \epsilon_j + (1-2\epsilon_j)\sum_{k=0}^{M-1} c_{k,j} h^{(j)}_{k|i}$$

and

$$\begin{aligned}\mathcal{D}_i &= \left\{\boldsymbol{u} \in \{0,1\}^N : d(\boldsymbol{u},\boldsymbol{c}_i) < \min_{0\le \ell \le M-1, \ell\ne i} d(\boldsymbol{u},\boldsymbol{c}_\ell)\right\} \\ &= \left\{\boldsymbol{u} \in \{0,1\}^N : \sum_{j=1}^{N}(u_j \oplus c_{i,j}) < \min_{0\le \ell \le M-1, \ell\ne i}\sum_{j=1}^{N}(u_j \oplus c_{\ell,j})\right\}\end{aligned}$$

is the decision partition for hypothesis H_i, are of the order $O(2^N NM)$. This order is much greater than that required by the criterion of (5). Simulations in the next section show that the DCFECC code obtained by minimizing (5) performs almost identically to the optimal DCFECC code that directly minimizes the fusion error. This justifies the feasibility of the use of criterion (5), in terms of both complexity and performance, for a large sensor network.

As far as the fault tolerance capability is concerned, another condition given in (12) should also be incorporated to constrain the minimum $d_{\min}$ required in the code search process. Without this minimum $d_{\min}$ constraint, the code that minimizes (5) may end up with limited fault-tolerance capability. Specifically, for the setting in Section VI-A, the 8×600 code matrix obtained by minimizing (5) without the minimum $d_{\min}$ constraint turns out to have a small $d_{\min} = 159$ (cf. Table I (see Section VI-C)), even though it has good pair-wise Hamming distances between adjacent hypotheses as $\min_{0\le \ell \le 7} d(\boldsymbol{c}_\ell, \boldsymbol{c}_{\ell+1}) = 263$. Since the

[2]The calculation of each $q_{i,j}$ given in (4) needs M exclusive-OR operations, $M+1$ multiplications, and M additions, and there are NM of $q_{i,j}$'s to determine. Hence, this step requires $O(NM^2)$ operations. It can be shown that with the availability of $\{q_{i,j}\}_{0\le i\le M-1, 1\le j\le N}$, the computation of (5) requires the same order of operations.

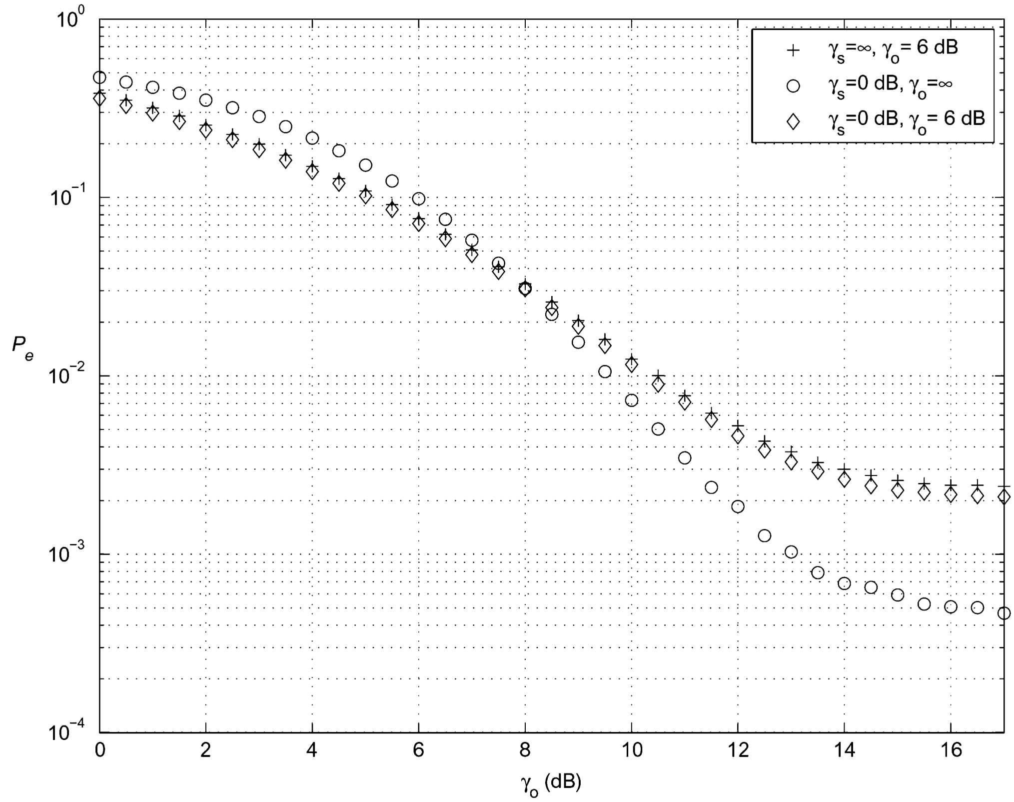

Fig. 4. Simulated performances of three 8×20 codes that minimize (9) without the minimum $d_{\min}$ constraint. The code search for these three codes are, respectively, performed at $(\gamma_s, \gamma_o) = (\infty,\ 6\ \text{dB})$, (0 dB, ∞) and (0 dB, 6 dB). The error probabilities in this figure are simulated at $\gamma_s = 0$ dB.

local classification error for distant hypotheses, such as H_0 and H_7, is very small when no sensors are faulty, it is unnecessary to provide a large pair-wise Hamming distance between distant hypotheses. However, as sensor faults can greatly enlarge the misclassification error for distant hypotheses, such code becomes less robust to sensor misbehaviors.

Hence, in what follows, we provide the search algorithm for the DCFECC code that minimizes (5) subject to the minimum $d_{\min}$ constraint in (12) for given system statistics, as well as sensor network size N and number of hypotheses M.

1. *Initialization.* For $1 \le j \le N$ and $0 \le \ell, i \le M-1$, assign[3]

$$h_{\ell|i}^{(j)} = \Pr\{y_j \in \Gamma_{\ell,j} | H_i\}$$

where

$$\Gamma_{\ell,j} \triangleq \left\{ y \in \mathcal{Y} : f_{\ell,j}(y) \ge \max_{0 \le i \le M-1, i \ne \ell} f_{i,j}(y) \right\}$$

and $f_{i,j}(y)$ is the probability density function of local observation at sensor j given hypothesis H_i is true.

2. For the previously assigned $\{h_{\ell|i}^{(j)}\}_{0 \le \ell, i \le M-1, 1 \le j \le N}$, find by simulated annealing algorithm the DCFECC code that minimizes (5) subject to the constraint of (12) for a target fault-tolerance capability $|\mathcal{F}|$ (or equivalently, a target sensor fault ratio defined as $|\mathcal{F}|/N$).

A final note in this section is that according to Lemma 3, the computational complexity of criterion (5) can be greatly reduced if the code search is applied to the simpler identical sensor system, where the quantities involved in computations are only the pair-wise Hamming distances.

[3]Here, the adopted local classification rules are not the system-wide globally optimal ones as those used in [18], but the locally optimal ones. Note that as the smallest size of WSNs considered in this paper, i.e., 20 sensors, is still much larger than the 10-sensor system considered in [18], global optimization of both local classification rules and the DCFECC code under minimum Hamming distance fusion becomes computational intractable. Nonetheless, with the tractable and easy-to-obtain locally optimal classification rules, the desired system performance, as well as the target fault-tolerance capability, can be achieved by deploying sufficient number of sensors, using the approach proposed in this work.

VI. Numerical and Simulation Results

In this section, we examine the performance and fault-tolerance capability of the DCFECC codes that are constructed through the code search algorithm of the previous section.

A. *Identical Sensor System*

Assume that each communication link employs antipodal transmission over an AWGN channel; hence, $\epsilon_j = \epsilon = \frac{1}{2}\text{erfc}(\sqrt{\gamma_s})$ for $1 \le j \le N$, where $\text{erfc}(\cdot)$ is the complementary error function, and γ_s is the signal-to-noise ratio of the communication link. In addition, each local observation y_j is assumed to be Gaussian distributed with mean ℓ and variance $1/\gamma_o$ given that hypothesis H_ℓ is true, where the signal-to-noise ratio for sensor observations, i.e., γ_o, is the square of the minimum difference in these Gaussian means,

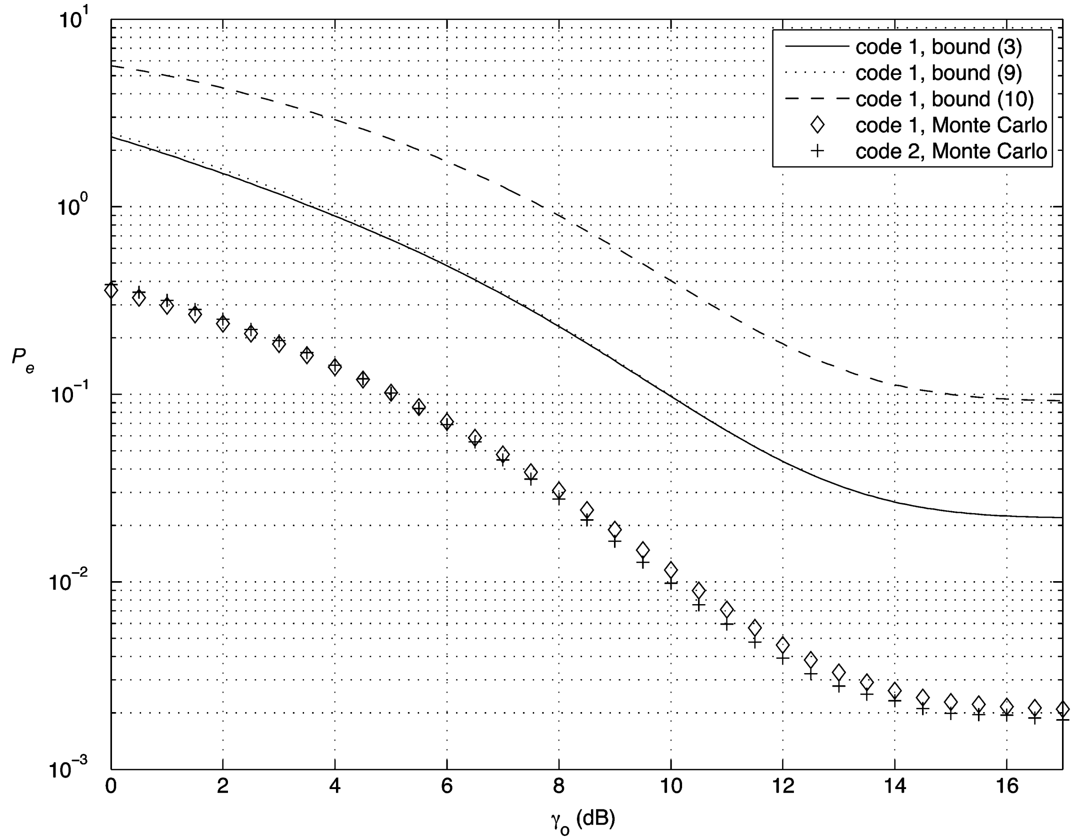

Fig. 5. Simulated performances of the two 8×20 codes that respectively minimizes (9) (code 1) and (15) (code 2) at $\gamma_s = 0$ dB. The code search is performed for $\gamma_s = 0$ dB and $\gamma_o = 6$ dB. Performance bounds for code 1 are also illustrated.

divided by the common observation variance. In this case, the code search criterion of (5) is reduced to (9).

Take $N = 20$ and $M = 8$. We first search for the DCFECC code that minimizes the minimum Hamming distance fusion error defined in (15), and compare it with the best code that minimizes (9) without the minimum $d_{\min}$ constraint. It needs to be pointed out that the best codes that respectively minimize (9) and (15) may be different for different target signal-to-noise ratios (see Fig. 4).[4] Here, we choose $\gamma_s = 0$ dB and $\gamma_o = 6$ dB as the target signal-to-noise ratios during the code search, which corresponds to $\epsilon \approx 0.07865$ and

$$\begin{bmatrix} h_{0|0} & h_{0|1} & h_{0|2} & \cdots & h_{0|7} \\ h_{1|0} & h_{1|1} & h_{1|2} & \cdots & h_{1|7} \\ h_{2|0} & h_{2|1} & h_{2|2} & \cdots & h_{2|7} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ h_{7|0} & h_{7|1} & h_{7|2} & \cdots & h_{7|7} \end{bmatrix}$$

$$= \begin{bmatrix} \rho_0 & \rho_{-1} & \rho_{-2} & \cdots & \rho_{-7} \\ \rho_1 - \rho_0 & \rho_0 - \rho_{-1} & \rho_{-1} - \rho_{-2} & \cdots & \rho_{-6} - \rho_{-7} \\ \rho_2 - \rho_1 & \rho_1 - \rho_0 & \rho_0 - \rho_{-1} & \cdots & \rho_{-5} - \rho_{-6} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{-7} & \rho_{-6} & \rho_{-5} & \cdots & \rho_0 \end{bmatrix}$$

$$\approx \begin{bmatrix} 0.8408 & 0.1592 & 0.0014 & \cdots & 0.0000 \\ 0.1578 & 0.6815 & 0.1578 & \cdots & 0.0000 \\ 0.0014 & 0.1578 & 0.6815 & \cdots & 0.0000 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0.0000 & 0.0000 & 0.0000 & \cdots & 0.8408 \end{bmatrix}$$

where $\rho_k = \Phi((k+0.5)\sqrt{\gamma_o})$, and $\Phi(\cdot)$ is the standard normal cumulative distribution function.

As shown in Fig. 5, code 1 and code 2 that are, respectively, obtained by minimizing (9) and (15), have almost indistinguishable performances even when the number of sensors is small. This is somewhat surprising because the performance bounds derived based on large deviations technique usually have a visible difference from the true performance for small sample size. This implies that the code search criterion (9) is not only much less complex than criterion (15), but can indeed yield a code that performs very close to the optimal minimum-fusion-error code. Furthermore, it can be observed that bounds (3) and (9) almost coincide with each other for all ranges of γ_o. In fact, it will be shown by subsequent simulations that the agreement between bounds (3) and (9) is not only true for small network size, but remains so as network size further increases. A final obser-

[4]One can observe from Fig. 4 that the three 8×20 codes obtained for different target signal-to-noise ratios have different performances. Their $(\lambda_{\max}, d_{\min})$ are $(-0.640353, 7)$, $(-0.523694, 10)$, and $(-0.57259, 7)$, respectively, for target signal-to-noise ratios $(\gamma_s, \gamma_o) = (\infty, 6\,\text{dB})$, $(0\,\text{dB}, \infty)$, and $(0\,\text{dB}, 6\,\text{dB})$. As anticipated, the code obtained at target $(\gamma_s, \gamma_o) = (0\,\text{dB}, \infty\text{dB})$ performs the best at high γ_o since its target signal-to-noise ratios are closest to the operational ones at this range, i.e., $\gamma_s =$ 0 dB and $\gamma_o \geq 8$ dB. Additionally, the code for target $(\gamma_s, \gamma_o) = (0$ dB,6 dB$)$ performs only slightly better than the code for target $(\gamma_s, \gamma_o) = (\infty$ dB,6 dB$)$, and both of them outperform the third code at low γ_o. From this figure, we can infer that the deviation of the target γ_o from the operational ones affects the resultant performance more than that of γ_s.

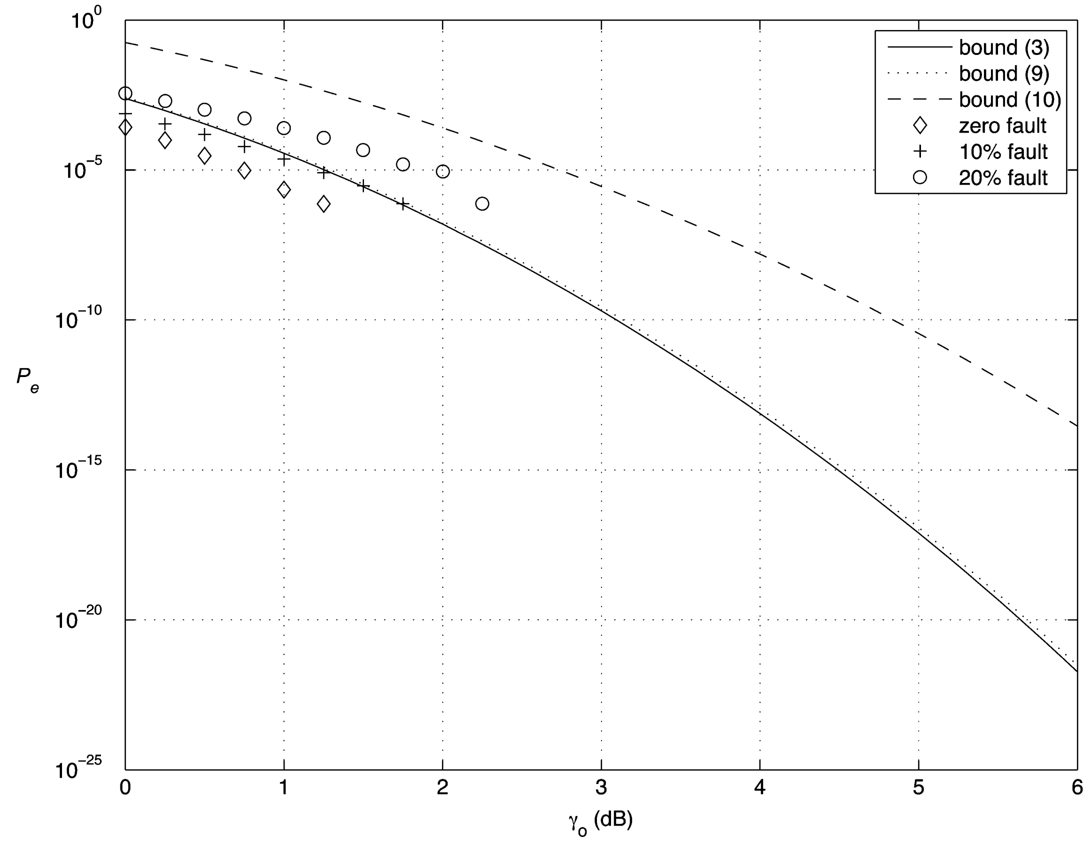

Fig. 6. Simulated performances and performance bounds of the 8×600 code that minimizes (9) subject to 10% sensor fault ratio at $\gamma_s = 0$ dB. The code search is performed for $\gamma_s = 0$ dB and $\gamma_o = 6$ dB. The type of the simulated sensor fault is stuck-at-one.

vation from Fig. 5 is that the fusion error of the obtained code achieves a floor value of 0.002 at large γ_o. Such an error-floor phenomenon is due to the existence of the communication link noise, and the floor value can be reduced when a larger γ_s is employed.

Next, we demonstrate that criterion (9) can be used for code search for a large sensor network. By taking $N = 600$ and $M = 8$, we search for the DCFECC code that minimizes the upper probability bound in (9) subject to 10% sensor fault ratio according to (12). Again, we choose $\gamma_o = 6$ dB and $\gamma_s = 0$ dB during the code search. The $\lambda_{\max}$ and $d_{\min}$ of the obtained code are, respectively, -0.548644 and 185. These two values, as expected, satisfy the minimum $d_{\min}$ constraint with 10% sensor fault ratio, i.e.,

$$\begin{aligned} d_{\min} &= 185 > -2|1-2\epsilon|\frac{|\mathcal{F}|}{\lambda_{\max}} \\ &= -2|1-2\times 0.07865|\frac{600\times 10\%}{(-0.548644)} = 184.316. \end{aligned}$$

The simulation results and performance bounds for this code are summarized in Figs. 6 and 7. Note that it is infeasible to perform the code search for minimum fusion error criterion defined in (15) for such a large sensor network.

From Fig. 6, we observe again that bound (3) is almost identical to bound (9), and is about 1.5 dB better than bound (10) at $P_e = 10^{-5}$. This indicates that it is not beneficial in terms of performance to replace the code search criterion (9) by the more complex (3), and adopting the simpler (10) as a new code search criterion may result in a code with degraded performance. Secondly, the simulated performance for the searched code with 10% stuck-at-one sensor fault ratio almost follows the curve of its search criterion (9) as we have anticipated, and is only 0.5 dB inferior to its performance without sensor faults at $P_e = 10^{-5}$. Also illustrated in this figure is that another 0.75-dB performance degradation will occur if the stuck-at-one sensor fault ratio increases up to 20%.

Observations similar to Fig. 6 can be made from Fig. 7 except that the performance degradation from zero fault to 10% sensor fault ratio is doubled when it is measured in γ_s rather in γ_o (specifically, 1 dB in Fig. 7, but only 0.5 dB in Fig. 6). The 3-dB performance difference between bound (9) and bound (10), when it is measured in γ_s, also becomes twice of the 1.5-dB difference for γ_o measure at $P_e = 10^{-5}$.

Finally, Fig. 8 compares the performances of four codes, respectively, obtained by minimizing bound (9) subject to zero sensor fault ratio, bound (9) subject to 10% sensor fault ratio, bound (10) subject to zero sensor fault ratio, and bound (10) subject to 10% sensor fault ratio. The target signal-to-noise ratios for these code searches are again $\gamma_s = 0$ dB and $\gamma_o = 6$ dB. We then gradually increase the number of faulty sensors to examine the robustness of these four codes. We observe that the code obtained by minimizing (9) subject to 10% sensor fault ratio performs the best when nearly 10% of sensors are faulty, and still remains the most robust when the number of faulty sensors is further increased. This result suggests again that bound (9) is a more suitable criterion to be minimized as far as a fault-tolerance WSN is concerned.

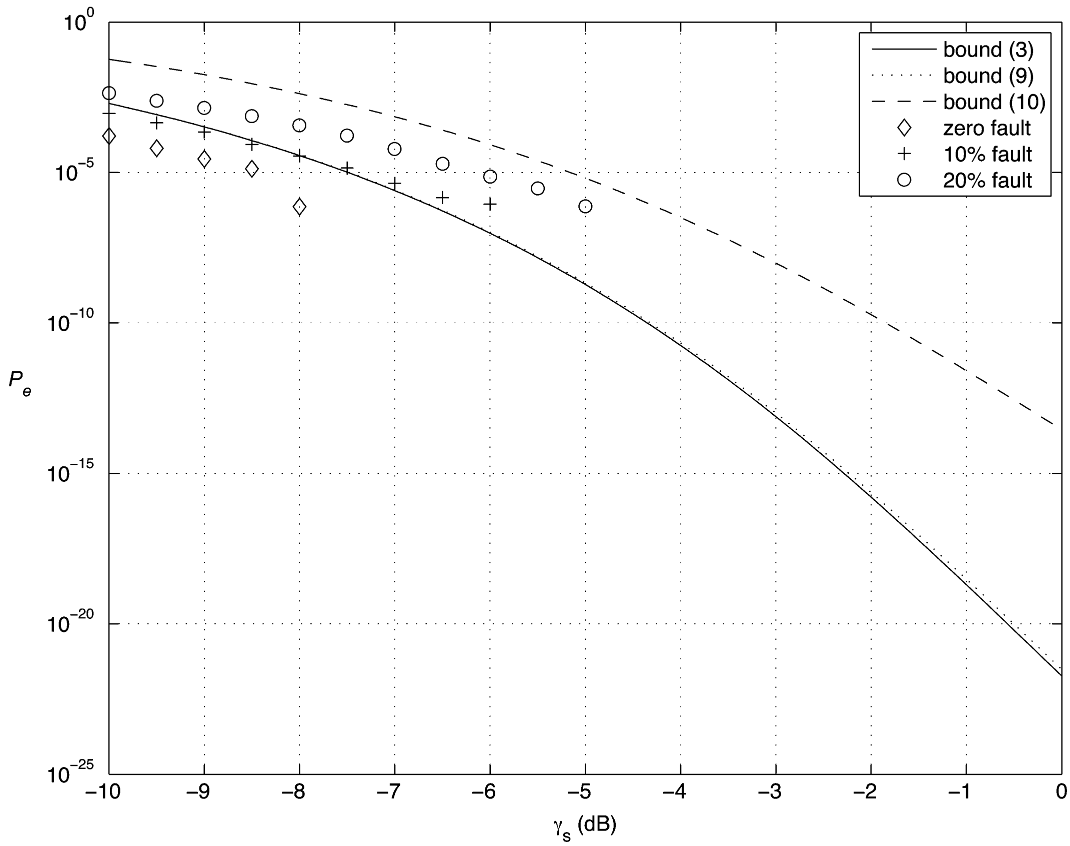

Fig. 7. Simulated performances and performance bounds of the 8×600 code that minimizes (9) subject to 10% sensor fault ratio at $\gamma_o = 6$ dB. The code search is carried out at $\gamma_s = 0$ dB and $\gamma_o = 6$ dB. The type of the simulated sensor fault is stuck-at-one.

B. Nonidentical Sensor System

In this simulation, the sensors are divided into three equal groups. Each sensor group forms an identical sensor subsystem as in the previous subsection. The signal-to-noise ratios for the three groups are, respectively, $(\gamma_o/2, 2\times\gamma_s)$, (γ_o, γ_s), and $(2\times \gamma_o, \gamma_s/2)$. In other words, some sensors have more accurate local classification but poor communication links, while some other sensors may be short in local classification accuracy but have less noisy communication links. A potential scenario for the above system configuration is that a group of sensors is closer to the target but farther from the fusion center, and another group of sensors is located distant from the target but is near the fusion center. Again, the code search is performed under $\gamma_o = 6$ dB and $\gamma_s = 0$ dB, and the target sensor fault ratio allowable is 10%. The $\lambda_{\max}$ and $d_{\min}$ of the obtained code are, respectively, -0.508553 and 250. The simulation results and performance bounds for this code are summarized in Figs. 9 and 10, and similar conclusions can be drawn as those for the identical sensor system. As a summary, in both figures, bound (5) coincides with bound (3), and the performance of the code found with the target sensor fault ratio follows the curve of its searched criterion (5). In addition, the performance degradation due to a decrement in γ_s is larger than that due to the same decrement in γ_o.

It should be pointed out that although bounds (3) and (5) (equivalently, bound (9) for the identical sensor system) coincide in the ranges in Figs. 7–10, they actually deviate from each other when γ_s and γ_o become much larger.[5] For example, when γ_s is large, bound (3) achieves a floor value of 1.56×10^{-34} in Fig. 7, while the ultimate floor value for bound (5) (namely, (9)) equals 5.19×10^{-34} in the same figure. A more clear difference in ultimate floor values can be obtained in Figs. 9 and 10, where bounds (3) and (5), respectively, achieve 1.34×10^{-58} and 2.38×10^{-53} at large γ_o in Fig. 9, and, respectively, approach 5.29×10^{-40} and 1.16×10^{-37} at large γ_s in Fig. 10. In addition, the ultimate floor values of bound (6) are 1.25×10^{-49}, 4.94×10^{-22}, 4.42×10^{-52}, and 1.08×10^{-29} for Figs. 6, 7, 9, and 10, respectively. Since these floor values indeed imply extremely small error performances that are of very minor interest in practice, we simply neglect the illustration of them in these figures.

We end this subsection by pointing out that similar observations can be made when the difference between the best and the worst sensors is further increased. As an example, by setting the signal-to-noise ratios, respectively, for the three groups as $(\gamma_o/4, 4\times\gamma_s)$, (γ_o, γ_s), and $(4\times\gamma_o, \gamma_s/4)$, searching the code under $\gamma_o = 8$ dB and $\gamma_s = 0$ dB subject to 10% sensor fault

[5]In Fig. 6, bounds (3) and (5) remain close even when γ_o is large. Since $q_{i,j}$ converges to ϵ_j as $\gamma_o \uparrow \infty$, it can be derived that bounds (3) and (5) (namely, (9)) actually achieve the same floor value of

$$\frac{1}{M}\sum_{i=0}^{M-1}\sum_{0\le\ell\le M-1,\ell\ne i}(1-(1-2\epsilon)^2)^{d(\boldsymbol{c}_\ell,\boldsymbol{c}_i)/2} = 1.79\times10^{-50}$$

when $\epsilon_j = \epsilon$ for all $1 \le j \le N$.

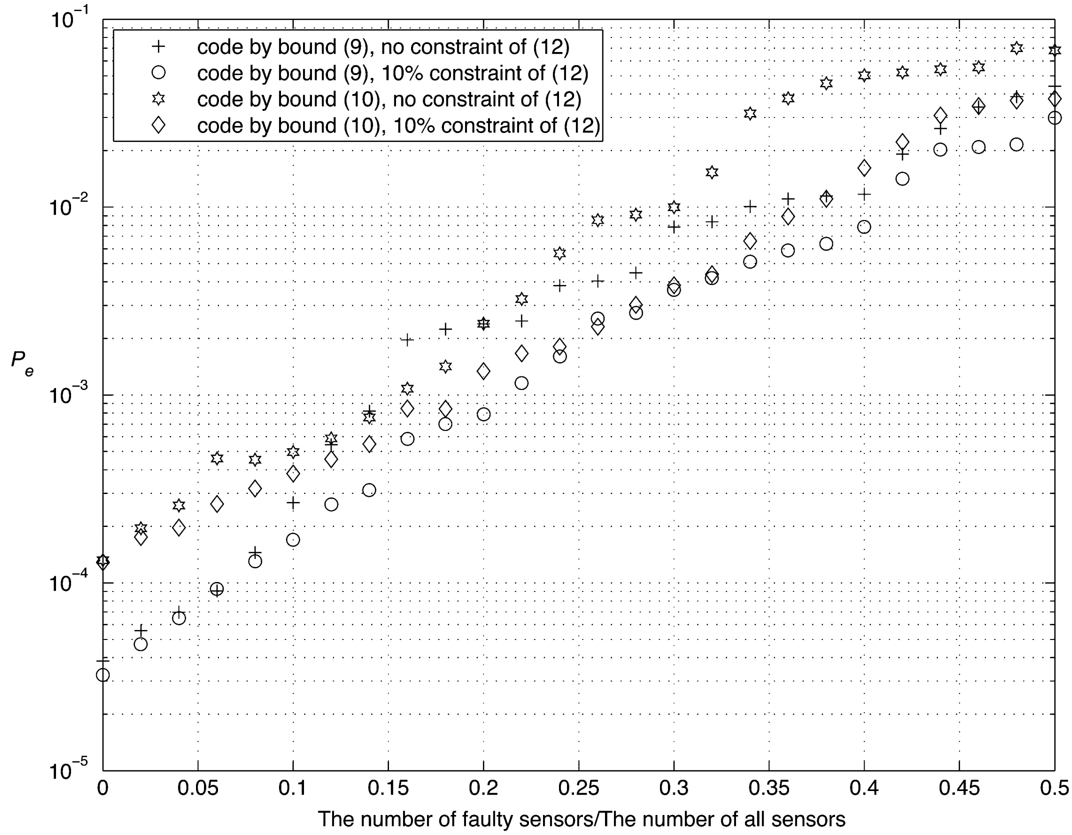

Fig. 8. Simulated performances of four 8×100 codes that minimizes bound (9), (10), (9) subject to 10% sensor fault ratio and bound (10) subject to 10% sensor fault ratio. The code search is performed at $\gamma_s = 0$ dB and $\gamma_o = 6$ dB. The type of the simulated sensor fault is stuck-at-one.

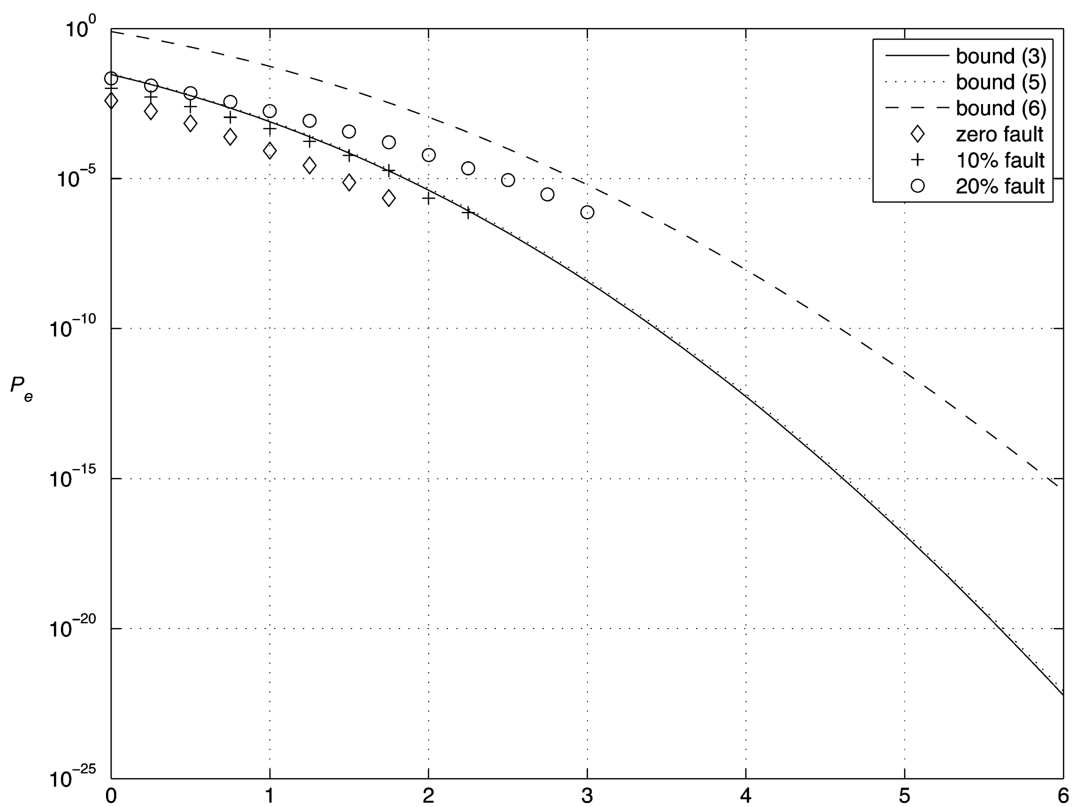

Fig. 9. Simulated performances and performance bounds of the 8×600 code that minimizes (5) subject to 10% sensor fault ratio at $\gamma_s = 0$ dB. The code search is performed at $\gamma_s = 0$ dB and $\gamma_o = 6$ dB. The type of the simulated sensor fault is stuck-at-one.

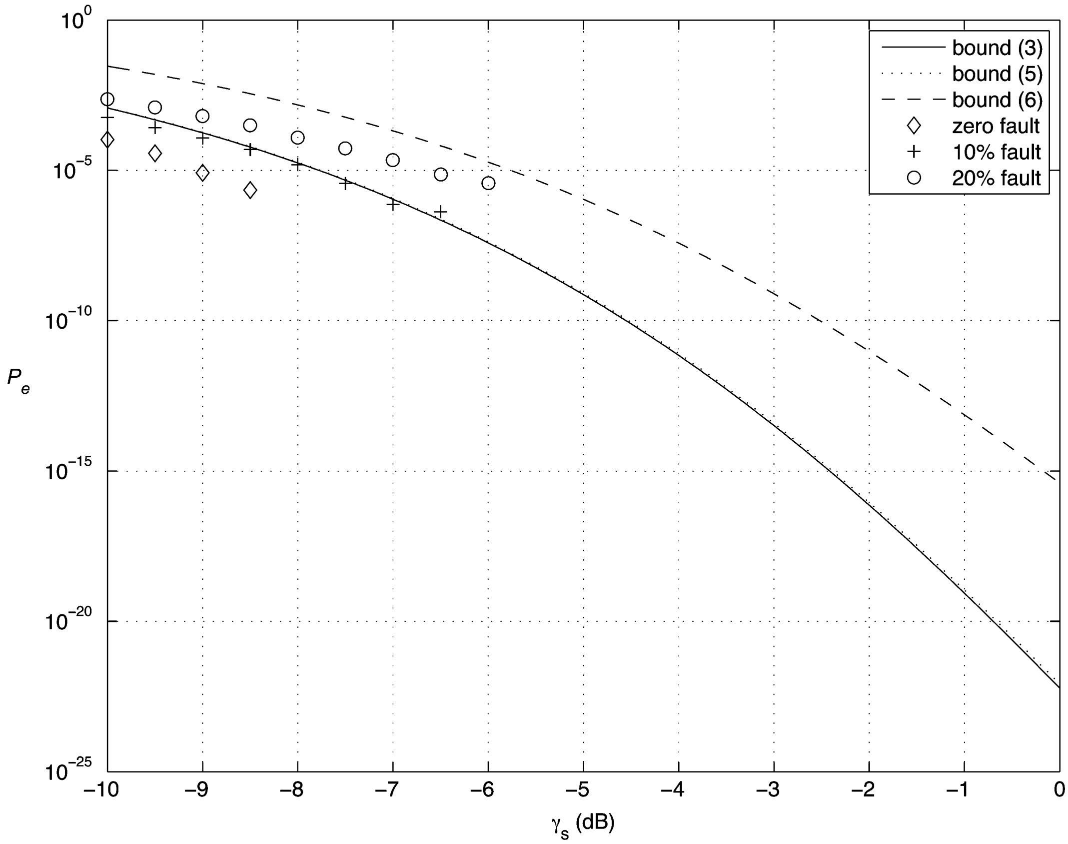

Fig. 10. Simulated performances and performance bounds of the 8×600 code that minimizes (5) subject to 10% sensor fault ratio at $\gamma_o = 6$ dB. The code search is performed at $\gamma_s = 0$ dB and $\gamma_o = 6$ dB. The type of the simulated sensor fault is stuck-at-one.

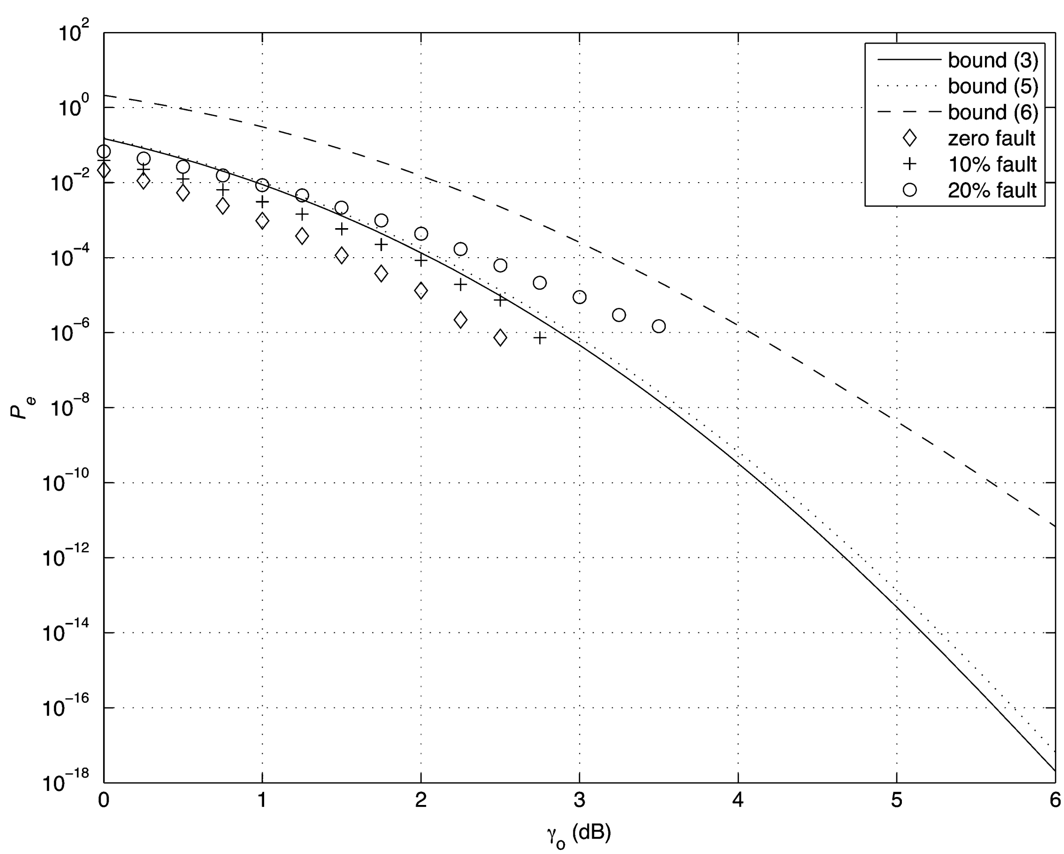

Fig. 11. Simulated performances and performance bounds of the 8×600 code that minimizes (5) subject to 10% sensor fault ratio at $\gamma_s = 0$ dB. The code search is performed at $\gamma_s = 0$ dB and $\gamma_o = 8$ dB. The type of the simulated sensor fault is stuck-at-one.

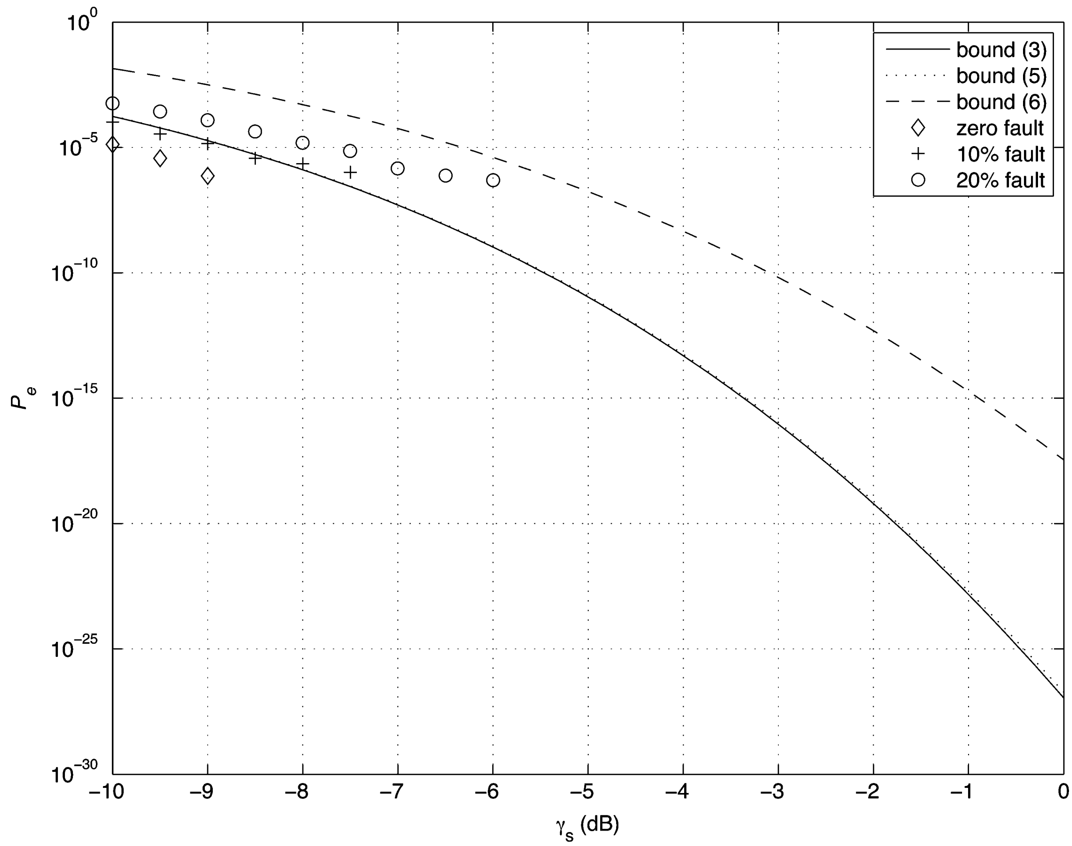

Fig. 12. Simulated performances and performance bounds of the 8×600 code that minimizes (5) subject to 10% sensor fault ratio at $\gamma_o = 8$ dB. The code search is performed at $\gamma_s = 0$ dB and $\gamma_o = 8$ dB. The type of the simulated sensor fault is stuck-at-one.

TABLE I
LIST OF EXPONENTS, $\lambda_{\max}$ AND $d_{\min}$ FOR THE BEST CODES THAT MINIMIZE (9) IN FIG. 13

N	$-\frac{1}{N}\log(P_e)$	$-\frac{1}{N}\log(\text{bound (9)})$	$-\frac{1}{N}\log(\text{bound (10)})$	$\lambda_{\max}$	$d_{\min}$	$\frac{d_{\min}}{N}$
50	0.112500	0.0685303	0.0215762	−0.547213	17	0.340000
100	0.103783	0.0785967	0.0314191	−0.560326	27	0.270000
150	0.103535	0.0802930	0.0392835	−0.552768	43	0.286667
200	——	0.0823383	0.0427901	−0.563429	55	0.275000
250	——	0.0833101	0.0430272	−0.565269	66	0.265000
300	——	0.0818457	0.0451263	−0.566490	80	0.266667
350	——	0.0840938	0.0458318	−0.563864	94	0.268571
400	——	0.0843196	0.0460971	−0.565061	106	0.265000
450	——	0.0843334	0.0469751	−0.563220	121	0.268889
500	——	0.0846376	0.0473064	−0.565257	133	0.266000
550	——	0.0845937	0.0479347	−0.563822	148	0.269091
600	——	0.0848210	0.0475872	−0.564467	159	0.265000
ave.	0.106606	0.0818094	0.0420796	−0.561766		0.275490

ratio yields Figs. 11 and 12. The $\lambda_{\max}$ and $d_{\min}$ of the obtained code are, respectively, -0.517076 and 271. It can then be observed the two figures that the results due to a larger difference in signal-to-noise ratios between the best and the worst sensors are actually very similar to those limited to 6-dB difference.

C. Fusion Error Versus Network Size and Number of Hypotheses

In this subsection, the best codes for the identical sensor system specified in Section VI-A are determined for different network sizes N under $M = 8$, $\gamma_o = 6$ dB, and $\gamma_s = 0$ dB. Since what we are mainly concerned with is the performance trend with respect to the network size, no minimum $d_{\min}$ constraint is set in this code search. Also simulated is the relation between resulting fusion error and number of hypothesis M under fixed $N = 100$. The results are summarized in Figs. 13 and 14.

We observe from Fig. 13 that the performance bounds for the best codes decrease exponentially with respect to the network size N, which hints that the true performances of the found

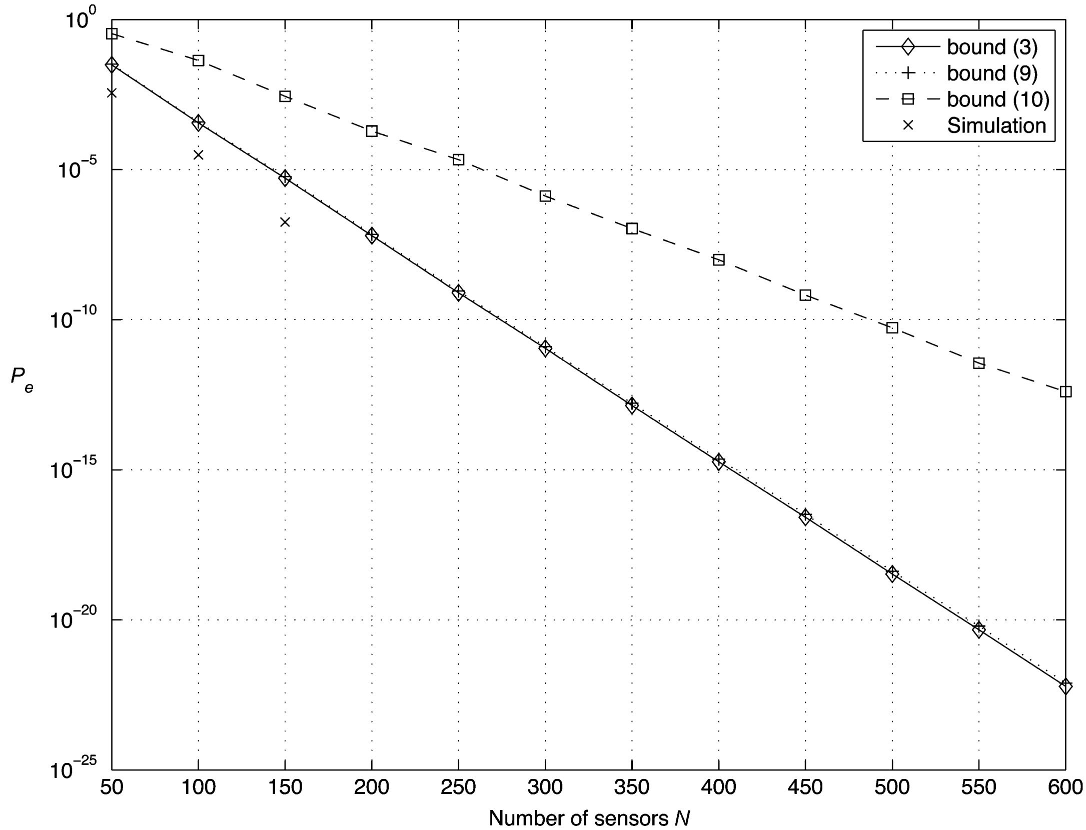

Fig. 13. Simulated performances and performance bounds for the family of best codes that minimize (9). The code search is performed at $M = 8, \gamma_s = 0$ dB, and $\gamma_o = 6$ dB.

TABLE II
LIST OF $\lambda_{\max}$ AND $d_{\min}$ FOR THE BEST CODES THAT MINIMIZE (9) IN FIG. 14

M	$\lambda_{\max}$	$d_{\min}$
10	−0.554042	25
20	−0.542517	30
30	−0.541723	29
40	−0.541113	31
50	−0.541057	31
60	−0.547502	32
70	−0.546961	32
80	−0.546335	36
90	−0.547075	34
100	−0.552319	35
110	−0.546452	34
120	−0.551985	34
2^7	−0.552439	35
2^8	−0.551318	35
2^9	−0.542934	29

DCFECC codes should at least follow the same trend of exponential decay. As shown in Table I, the exponents for the simulated performances, bound (9) and bound (10) are, respectively, $0.1066, 0.0818$, and 0.0421 on an average. Also listed in Table I are the $\lambda_{\max}$ and $d_{\min}$ for these searched codes. The list indicates that $\lambda_{\max}$ for the best code is almost independent of the network size, and remains around -0.56 for most N. In addition, $d_{\min}$ increases linearly with respect to N.

Fig. 14 summarizes the simulated relationship between fusion error and the number of hypotheses. Recall that the code rate of an $M \times N$ code matrix is conventionally defined as $R = (1/N)\log_2(M)$. Thus, the code rates of the DCFECC codes simulated previously are prohibitively small when they are compared with the code rates of the traditional error-correcting codes. Since the traditional error-correcting codes with *larger* code rates (or equivalently, larger M) can still yield a good $d_{\min}$ (e.g., there exists rate-$(8/127)$ Bose–Chaudhuri–Hocquenghem (BCH) code with $d_{\min} = 63$ [15, p. 438]), it is not surprising to obtain that at $N = 100$, the fusion error only mildly increases with respect to M when $M \leq 120$. Even with a scale of exponentially increasing M up to 2^9, the growth rate of the fusion errors only sightly increases. It can be observed from Table II that the $d_{\min}$ of the best codes that minimize (9) does not tend to decrease in M for $M \leq 120$ (as it should do for very large M). Again, $\lambda_{\max}$ remains almost constant for all M, and has the average around -0.56. This result indicates that the number of hypotheses for the DCFECC codes can actually be of an exponential order in the number of sensors (specifically, $M = 2^{RN}$).

VII. CONCLUSION

The envisioned use of sensor networks for fault-tolerant classification applications in a harsh environment calls for the de-

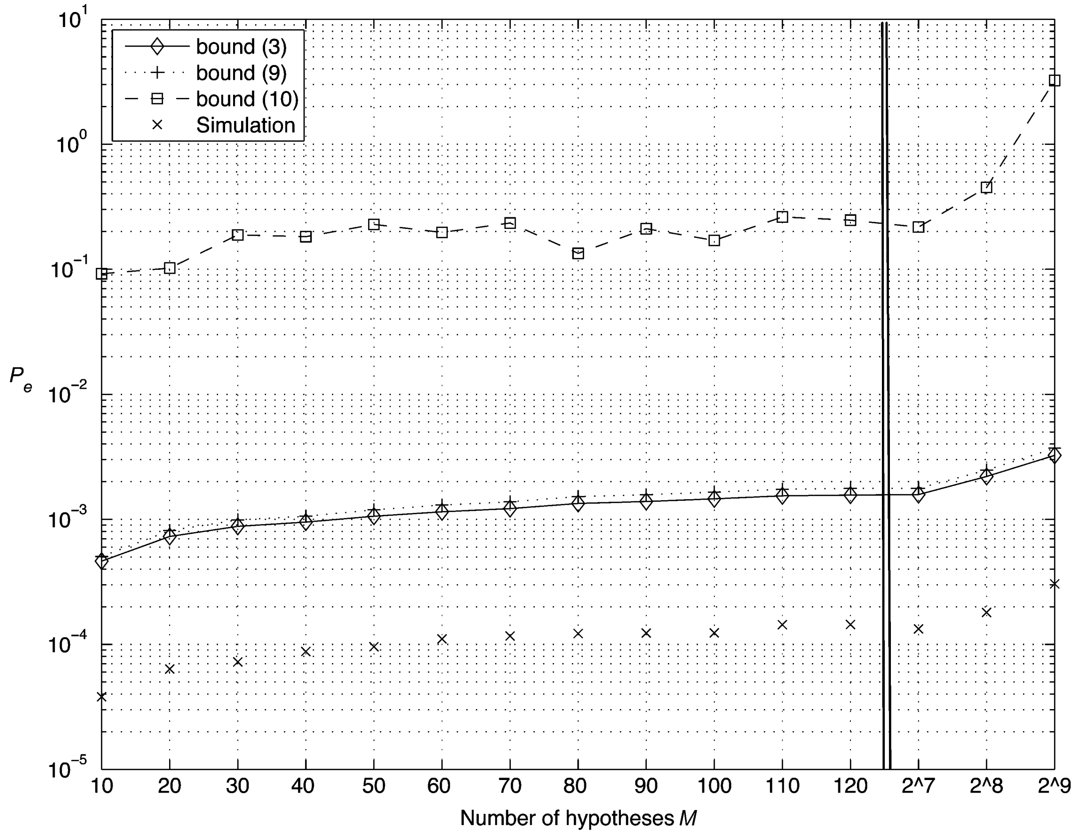

Fig. 14. Simulated performances and performance bounds for the family of best codes that minimize (9). The code search is performed at $N = 100, \gamma_s = 0$ dB, and $\gamma_o = 6$ dB.

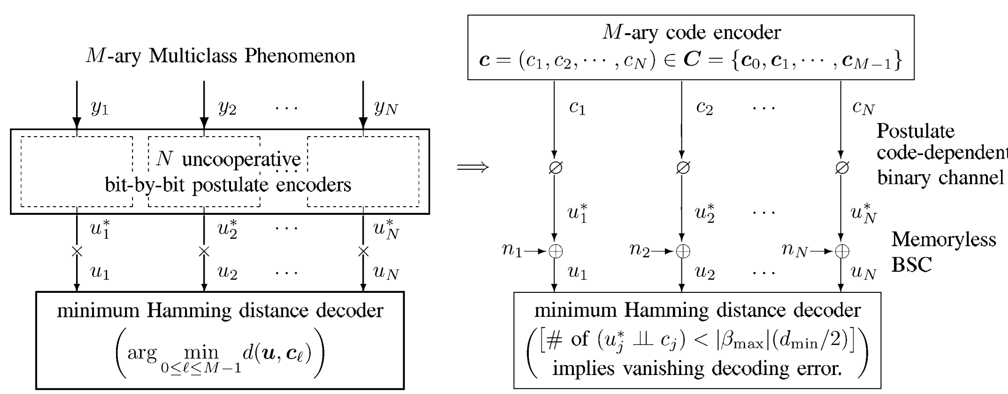

Fig. 15. Equivalent serial-connected binary channel model specifically for wireless sensor networks. The channel noise $\{n_j\}_{j=1}^N$ for memoryless BSC is independent and identically distributed with $\Pr\{n_j = 1\} = \epsilon$. The transition probability $\Pr\{u_j^* | c_j\}$ is given by (16), and is a function of the codebook $\boldsymbol{C}$.

sign of collaborative classification using coding theory, e.g., the DCFECC approach with minimum Hamming distance fusion. In this paper, we have analytically characterized its performance in both fault-free and faulty situations. Also, characterized is the fault-tolerance capability of a DCFECC code, which is subsequently used together with the newly proposed code search criterion to find the code with the desired fault-tolerant robustness. Our results indicate, as anticipated, the strong relation between the fault-tolerance capability and the pair-wise Hamming distances of a DCFECC code. Although we mostly emphasize that our code construction approach is specifically suitable for networks with a large number of sensors owing to its simplicity, the simulation result shown in Fig. 5 implies its suitability in terms of performance for networks with only tens of sensors.

The coding problem considered in this paper can actually be transformed into one for the memoryless binary-symmetric channel (BSC) with *unreliable bit-by-bit postulate encoders* as shown in Fig. 15, if the link noises have common marginal distribution. We can further consider the memoryless BSC channel with unreliable bitwise postulate encoders as a serial connection of two binary channels, in which the first channel suffers *code-dependent* noises that give

$$\Pr(u_j^* \,|\, c_j) = \frac{\sum_{i=0}^{M-1} \left\{ [1 - (c_j \oplus c_{i,j})] \sum_{k=0}^{M-1} [1 - (u_j^* \oplus c_{k,j})] h_{k|i}^{(j)} \right\}}{\sum_{i=0}^{M-1} [1 - (c_j \oplus c_{i,j})]} \tag{16}$$

and the second channel is the memoryless BSC channel. The case of sensor faults under the equivalent channel model becomes that of u_j^* turning independent of c_j (and hence, code independent) without notifying the fusion center. Our results then indicate that the constraint that the number of code-independent bits in $\boldsymbol{u}^*$ (i.e., the number of faulty sensors) is less than $|\beta_{\max}| \times (d_{\min}/2)$ is sufficient to guarantee a vanishing decoding error for such a serially connected binary channel. This bound is derived based on the pessimistic view when both indices of faulty sensors and sensor fault types are unknown to the fusion center, or equivalently, the decoder is aware of neither the index of every *faulty* bit u_j^* nor its resultant *code-independent* distribution. In the extreme case that $\boldsymbol{u}^*$ and $\boldsymbol{c}$ are completely dependent, which should occur when $h_{k|i}^{(j)} = 1$ for every $0 \le k = i \le M-1$, the constraint reduces to the conventional $|\mathcal{F}| < d_{\min}/2$ for coding techniques since $\beta_{\max} = -1$ as we have anticipated. This observation hints that in a channel suffering from code-dependent noises, a code that makes $\boldsymbol{u}^*$ (channel output) and $\boldsymbol{c}$ (channel input) more "dependent" (and thus, the channel output has more information about the input) is expected to be a better and more robust code, which is exactly the underlying concept behind the Shannon baptized "channel capacity." It would be interesting to research along this line, and determine the capacity of the postulate code-dependent channels.

Appendix

A. Proof of Lemma 1

1) Property of $I_m(x)$: $I_m(x)$ is convex since it is the pointwise supremum of a collection of affine functions. Also, $I_m(x) \ge 0$ because $[\theta x - \varphi_m(\theta)] = 0$ for $\theta = 0$. By Jensen's inequality

$$\begin{aligned}\varphi_m(\theta) &= \frac{1}{m}\log E[\exp\{\theta(Z_1+\cdots+Z_m)\}]\\ &\ge \frac{1}{m}\log\left(\exp\{\theta\cdot E[Z_1+\cdots+Z_m]\}\right)\\ &= \theta\cdot\frac{1}{m}E[Z_1+\cdots+Z_m] = \theta\lambda_m\end{aligned}$$

which implies $\theta\lambda_m - \varphi_m(\theta) \le 0$. Hence, $I_m(x)$ gives its minimum value 0 at $x = \lambda_m$.

2) Support Line of $I_m(x)$: Let $[\theta_0 x - \varphi_m(\theta_0)]$ be the support line of the convex $I_m(x)$, which passes through the point $(0, I_m(0))$. The convexity of $I_m(x)$ and $\lambda_m < 0$ imply that $\theta_0 \ge 0$, and

$$[0,\infty) \subset \{x \in \Re : \theta_0 x - \varphi_m(\theta_0) \ge I_m(0)\}.$$

3) Probability Bound:

$$\begin{aligned}&\Pr\left[Z_1+\cdots+Z_m \ge 0\right]\\ &\quad\le \Pr\left[\frac{Z_1+\cdots+Z_m}{m} \in \{x\in\Re : \theta_0 x - \varphi_m(\theta_0) \ge I_m(0)\}\right]\\ &\quad= \Pr\left[\theta_0\frac{Z_1+\cdots+Z_m}{m} - \varphi_m(\theta_0) \ge I_m(0)\right]\\ &\quad= \Pr[\exp\{\theta_0(Z_1+\cdots+Z_m)\}\\ &\qquad\ge \exp\{m\varphi_m(\theta_0) + mI_m(0)\}]\\ &\quad\le \frac{E[\exp\{\theta_0(Z_1+\cdots+Z_m)\}]}{\exp\{m\varphi_m(\theta_0)+mI_m(0)\}}\\ &\quad= e^{-m\cdot I_m(0)}.\end{aligned}$$

B. Proof of Lemma 2

Let $\bar{q}_m = (1/m)\sum_{i=1}^m q_j$, and note that $\lambda_m = 2\bar{q}_m - 1$. Therefore, the assumption of the lemma is equivalent to $\bar{q}_m < 1/2$.

The validity of the lemma for $0 < \bar{q}_m < 1/2$ can be proved by Jensen's inequality in terms of the upper bound in (1) as follows:

$$\begin{aligned}&\inf_{\theta\ge 0}\exp\left\{\sum_{j=1}^m \log\left(q_j e^{\theta} + (1-q_j)e^{-\theta}\right)\right\}\\ &\quad= \inf_{\theta\ge 0}\exp\left\{m\left(\sum_{j=1}^m \frac{1}{m}\log\left(q_j e^{\theta}+(1-q_j)e^{-\theta}\right)\right)\right\}\\ &\quad\le \inf_{\theta\ge 0}\exp\left\{m\cdot\log\left(\bar{q}_m e^{\theta} + (1-\bar{q}_m)e^{-\theta}\right)\right\}\\ &\quad= (4\bar{q}_m(1-\bar{q}_m))^{m/2}\end{aligned}$$

where the last equality takes the optimizer

$$\theta^* = \log\sqrt{(1-\bar{q}_m)/\bar{q}_m} > 0, \quad \text{for } 0 < \bar{q}_m < 1/2.$$

In case $\bar{q}_m = 0$, we have $(4\bar{q}_m(1-\bar{q}_m))^{m/2} = 0$ and

$$\begin{aligned}&\inf_{\theta\ge 0}\exp\left\{\sum_{j=1}^m \log\left(q_j e^{\theta}+(1-q_j)e^{-\theta}\right)\right\}\\ &\quad\le \inf_{\theta\ge 0}\exp\left\{m\log\left(\bar{q}_m e^{\theta}+(1-\bar{q}_m)e^{-\theta}\right)\right\}\\ &\quad= \inf_{\theta\ge 0}\exp\{-m\theta\} = 0.\end{aligned}$$

C. Proof of Lemma 3

Under the assumption of an identical sensor system

$$q_{i,j} = \epsilon + (1-2\epsilon)\sum_{k=0}^{M-1}(c_{i,j}\oplus c_{k,j})h_{k|i}.$$

Thus

$$\begin{aligned}&\sum_{j=1}^N (c_{\ell,j}\oplus c_{i,j})q_{i,j}\\ &\quad= \sum_{j=1}^N (c_{\ell,j}\oplus c_{i,j})\left(\epsilon + (1-2\epsilon)\sum_{k=0}^{M-1}(c_{i,j}\oplus c_{k,j})h_{k|i}\right)\\ &\quad= \epsilon\cdot d(\boldsymbol{c}_\ell,\boldsymbol{c}_i) + (1-2\epsilon)\sum_{k=0}^{M-1}h_{k|i}\sum_{j=1}^N (c_{\ell,j}\oplus c_{i,j})(c_{i,j}\oplus c_{k,j})\\ &\quad= \epsilon\cdot d(\boldsymbol{c}_\ell,\boldsymbol{c}_i) + (1-2\epsilon)\\ &\qquad\times\sum_{k=0}^{M-1}h_{k|i}\frac{d(\boldsymbol{c}_\ell,\boldsymbol{c}_i)+d(\boldsymbol{c}_i,\boldsymbol{c}_k)-d(\boldsymbol{c}_\ell,\boldsymbol{c}_k)}{2}\\ &\quad= \frac{1}{2}d(\boldsymbol{c}_\ell,\boldsymbol{c}_i) + \frac{(1-2\epsilon)}{2}\sum_{k=0}^{M-1}h_{k|i}\left(d(\boldsymbol{c}_i,\boldsymbol{c}_k) - d(\boldsymbol{c}_\ell,\boldsymbol{c}_k)\right).\end{aligned}$$

Accordingly

$$\begin{aligned}\frac{1}{d(\boldsymbol{c}_\ell,\boldsymbol{c}_i)}&\sum_{j=1}^{N}(c_{\ell,j}\oplus c_{i,j})(2q_{i,j}-1)\\&=\frac{2}{d(\boldsymbol{c}_\ell,\boldsymbol{c}_i)}\left(\frac{1}{2}d(\boldsymbol{c}_\ell,\boldsymbol{c}_i)\right.\\&\quad\left.+\frac{(1-2\epsilon)}{2}\sum_{k=0}^{M-1}h_{k|i}(d(\boldsymbol{c}_i,\boldsymbol{c}_k)-d(\boldsymbol{c}_\ell,\boldsymbol{c}_k))\right)-1\\&=\frac{(1-2\epsilon)}{d(\boldsymbol{c}_\ell,\boldsymbol{c}_i)}\sum_{k=0}^{M-1}h_{k|i}\left[d(\boldsymbol{c}_i,\boldsymbol{c}_k)-d(\boldsymbol{c}_\ell,\boldsymbol{c}_k)\right]\end{aligned}$$

and the lemma is verified.

REFERENCES

[1] 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.

[2] S. A. Aldosari and J. M. F. Moura, "Detection in decentralized sensor networks," in *Proc. IEEE Int. Conf. Accoustics, Speech, and Signal Processing*, Montreal, QC, Canada, May 2004, pp. 277–280.

[3] J.-F. Chamberland and V. V. Veeravalli, "Decentralized detection in sensor networks," *IEEE Trans. Signal Process.*, vol. 51, no. 2, pp. 735–744, Feb. 2003.

[4] J.-F. Chamberland and V. V. Veeravalli, "Asymptotic results for decentralized detection in power constrained wireless sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 22, no. 6, pp. 1007–1015, Aug. 2004.

[5] P.-N. Chen, "Generalization of Gartner-Ellis theorem," *IEEE Trans. Inf. Theory*, vol. 46, no. 7, pp. 2752–2759, Nov. 2000.

[6] P.-N. Chen and A. Papamarcou, "New asymptotic results in parallel distributed detection," *IEEE Trans. Inf. Theory*, vol. 39, no. 6, pp. 1847–1863, Nov. 1993.

[7] P.-N. Chen, T.-Y. Wang, Y. S. Han, P. K. Varshney, and C. Yao, "Asymptotic performance analysis for minimum-hamming-distance fusion," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, Philidaphia, PA, May 2005, pp. 865–868.

[8] L. Dan, K. D. Wong, H. H. Yu, and A. M. Sayeed, "Detection, classification, and tracking of targets," *IEEE Signal Process. Mag.*, vol. 19, no. 2, pp. 17–29, Mar. 2002.

[9] A. D'Costa, V. Ramachandran, and A. M. Sayeed, "Distributed classification of Gaussian spaceVtime sources in wireless sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 22, no. 6, pp. 1026–1036, Aug. 2004.

[10] A. D'Costa and A. M. Sayeed, "Data versus decision fusion in sensor networks," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, Hong Kong, China, Apr. 2003, pp. 832–835.

[11] A. D'Costa and A. M. Sayeed, "Data versus decision fusion for classification in sensor networks," in *Proc. 6th Int. Conf. Information Fusion*, Cains, Australia, Jul. 2003, pp. 889–894.

[12] S. K. Jayaweera, "Large system decentralized detection performance under communication constraints," *IEEE Commun. Lett.*, vol. 9, no. 9, pp. 769–771, Sep. 2005.

[13] S. Lin and D. J. Costello Jr., *Error Control Coding: Fundamentals and Applications*. Englewood Cliffs, NJ: Prentice-Hall, 1983.

[14] Z.-Q. Luo, "An isotropic universal decentralized estimation scheme for a bandwidth constrained ad hoc sensor network," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 4, pp. 735–744, Apr. 2005.

[15] J. G. Proakis, *Digital Communications*, 4th ed. New York: McGraw-Hill, 2001.

[16] J. N. Tsitsiklis, "Decentralized detection by a large number of sensors," *Math. Conrol, Signals, Syst.*, vol. 1, no. 2, pp. 167–182, 1988.

[17] H. Wang, J. Elson, L. Girod, D. Estrin, and K. Yao, "Target classification and localization in habitat monitoring," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, Hong Kong, China, Apr. 2003, pp. 844–847.

[18] T.-Y. Wang, Y. S. Han, P. K. Varshney, and P.-N. Chen, "Distributed fault-tolerant classification in wireless sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 4, pp. 724–734, Apr. 2005.

[19] Y. Yuan and M. Kam, "Distributed decision fusion with a random-access channel for sensor network applications," *IEEE Trans. Intrumem. Meas.*, vol. 53, no. 4, pp. 1339–1344, Aug. 2004.