

Chernoff Information-Based Optimization of Sensor Networks for Distributed Detection

Gernot Fabeck, Rudolf Mathar
Institute for Theoretical Information Technology
RWTH Aachen University
D-52056 Aachen, Germany
Email: {fabeck, mathar}@ti.rwth-aachen.de

Abstract—This paper addresses the scalable optimization of sensor networks for distributed detection applications. In the general case, the jointly optimum solution for the local sensor decision rules and the fusion rule is extremely difficult to obtain and does not scale with the number of sensors. In this paper, we consider optimization of distributed detection systems based on a local metric for sensor detection performance. Derived from the asymptotic error exponents in binary hypothesis testing, the Chernoff information emerges as an appropriate metric for sensor detection quality. By locally maximizing the Chernoff information at each sensor and thus decoupling the optimization problem, scalable solutions are obtained which are also robust with respect to the underlying prior probabilities. By considering the problem of detecting a deterministic signal in the presence of Gaussian noise, a detailed numerical study illustrates the feasibility of the proposed approach.

I. INTRODUCTION

Distributed detection is one of the primary applications of wireless sensor networks and is often the first step in an overall sensing process [1]–[3]. The nodes in a sensor network typically operate on limited energy budgets and are consequently subject to communication constraints. This recommends compression of observations at the sensors and transmission of quantized observations or local decisions. In the parallel fusion network, the sensor nodes process their observations independently and make preliminary decisions about the state of the observed environment, e.g., absence or presence of a target. The sensors transmit the local decisions to a fusion center that combines the received decisions and computes the final detection result. Since the transmission channels between the wireless sensors and the fusion center are subject to noise and interference, it might also be necessary to take wireless channel conditions into account [4].

The main problem is to design the local sensor decision rules and the fusion rule with respect to an overall performance criterion, e.g., minimum probability of error. In the general case, the jointly optimum solution for the local decision rules at the sensors and the fusion rule is very difficult to obtain and does not scale with the number of sensors. Global optimization of distributed detection systems was first investigated by Reibman and Nolte [5]. They consider simultaneous optimization of binary local detectors and the fusion rule under the constraints of identical local sensor decision rules and restrictions on the employed fusion rule. Numerical algorithms

that find person-by-person optimal local sensor decision rules are presented in [6] and [7]. In [8], the authors use an iterated combination of a genetic algorithm for optimizing the fusion rule and a gradient-based algorithm to optimize the decision thresholds of the local detectors. All of the above authors assume independent and identically distributed observations at the local sensors and the joint optimization is done only for sensor networks with a very low number of sensors, e.g., in [8] the number of sensors varies between 2 and 8.

Another interesting approach to the optimization of distributed detection systems is presented in [9]. The authors decrease the computational complexity of the original optimization problem by using distributional distances as objective function instead of the original minimum probability of error criterion. The local sensor decision rules are obtained by solving a system of coupled nonlinear equations which in general has multiple solutions. Besides the fact that their approach is only applicable to binary quantization at the sensors, the main drawback lies in the coupled optimization problem which again restricts the number of sensors considered in the optimization to a maximum of 15.

In this paper, we consider the optimization of distributed detection systems based on a local metric for sensor detection performance. An appropriate metric is derived from the asymptotic error exponents in binary hypothesis testing and is given by the Chernoff information between probability distributions. By locally maximizing the Chernoff information between the probability vectors of quantization probabilities at each sensor, a decoupling of the optimization problem is obtained. Channel state information at the sensors might also be used in the optimization procedure. After the local sensor decision rules have been determined, the optimal channel-aware fusion rule can be derived. As the numerical results show, the presented approach enables scalable design of near-optimal distributed detection systems and is also applicable to realistic scenarios with nonhomogenous sensing conditions.

The remainder of this paper is organized as follows. In Section II, the problem of distributed detection with M -ary quantization, noisy channels, and soft decision fusion is stated. The Chernoff information-based optimization procedure is motivated and presented in Section III. In Section IV, a detailed numerical analysis of the proposed approach is given. Finally, we conclude in Section V.

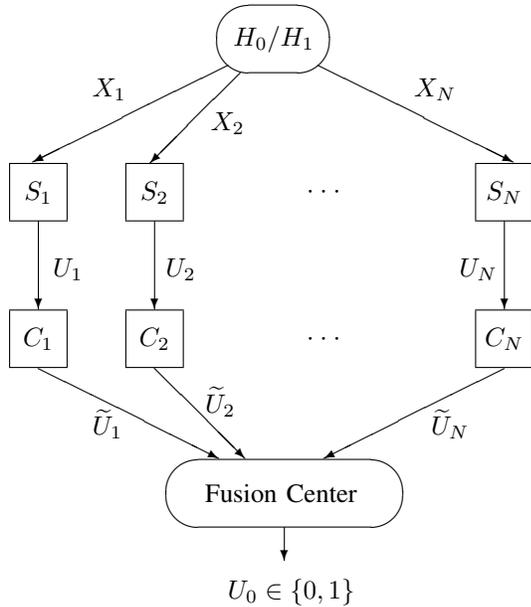


Fig. 1. Parallel fusion network with noisy channels.

II. DISTRIBUTED DETECTION

The problem of distributed detection in parallel fusion networks with M -ary quantization at the local sensors, noisy channels and soft decision fusion at the fusion center can be stated as follows (see Fig. 1). We consider a binary hypothesis testing problem with hypotheses H_0 and H_1 indicating the state of the monitored environment. The associated prior probabilities are $\pi_0 = P(H_0)$ and $\pi_1 = P(H_1)$. In order to detect the true state of nature, a network of N sensors S_1, \dots, S_N obtains random observations

$$(X_1, \dots, X_N)' \in \mathcal{X}_1 \times \dots \times \mathcal{X}_N, \quad (1)$$

which are generated according to either H_0 or H_1 . The random observations X_1, \dots, X_N are assumed to be conditionally independent across sensors given the underlying hypothesis, i.e., the joint conditional probability density function of all the observations factorizes according to

$$f(x_1, \dots, x_N | H_k) = \prod_{j=1}^N f_j(x_j | H_k), \quad k = 0, 1. \quad (2)$$

According to the distributed nature of the problem, the sensors process their respective observations X_j independently by forming local decisions

$$U_j = \delta_j(X_j), \quad j = 1, \dots, N. \quad (3)$$

Thus, the local decision U_j of sensor S_j does only depend on its own observation X_j and not on the observations of the other sensors.

A. Local sensor decision rules

In the general case of M -ary quantization at the local sensors, the local sensor decision rules δ_j are mappings

$$\delta_j: \mathcal{X}_j \rightarrow \{1, \dots, M\}, \quad j = 1, \dots, N. \quad (4)$$

Warren and Willett have shown that the sensor decision rules leading to jointly optimal configurations under the minimum probability of error criterion are monotone likelihood ratio partitions of the sensor observation spaces $\mathcal{X}_1, \dots, \mathcal{X}_N$, provided that the observations are conditionally independent across sensors [10]. Hence, it is necessary only to consider sensor decision rules δ_j that can be parameterized by a set of real quantization thresholds $\tau_{j_1}, \dots, \tau_{j_{M-1}}$, where $\tau_{j_0} = -\infty$, $\tau_{j_M} = \infty$, and $\tau_{j_k} \leq \tau_{j_{k+1}}$. In this way, each sensor S_j is characterized by the conditional quantization probabilities

$$\alpha_{j_k} = P(U_j = k | H_0) = P(\tau_{j_{k-1}} < L_j \leq \tau_{j_k} | H_0), \quad (5)$$

$$\beta_{j_k} = P(U_j = k | H_1) = P(\tau_{j_{k-1}} < L_j \leq \tau_{j_k} | H_1), \quad (6)$$

where $L_j = \log(f_j(X_j | H_1) / f_j(X_j | H_0))$ is the local log-likelihood ratio of observation X_j . The probability vectors $\alpha_j = (\alpha_{j_1}, \dots, \alpha_{j_M})'$ and $\beta_j = (\beta_{j_1}, \dots, \beta_{j_M})'$ are computable given the local observation statistics $f_j(\cdot | H_k)$ and the quantization thresholds $\tau_{j_1}, \dots, \tau_{j_{M-1}}$ for each $j = 1, \dots, N$.

B. Transmission of local decisions

Upon local decision-making, the sensor nodes transmit their local decisions

$$(U_1, \dots, U_N)' \in \{1, \dots, M\}^N \quad (7)$$

to the fusion center in order to perform decision combining. We model the communication link C_j between sensor S_j and the fusion center by a discrete noisy channel with transition matrix \mathbf{T}_j . The channel transition matrix $\mathbf{T}_j = (T_{kl}^{(j)})_{1 \leq k, l \leq M}$ is an $M \times M$ matrix with the kl th entry defined as

$$T_{kl}^{(j)} = P(\tilde{U}_j = k | U_j = l), \quad k, l \in \{1, \dots, M\}, \quad (8)$$

where $\sum_{k=1}^M T_{kl}^{(j)} = 1$ for any $l \in \{1, \dots, M\}$. Because of the noisy channels, the fusion center receives a vector of potentially corrupted decisions

$$(\tilde{U}_1, \dots, \tilde{U}_N)' \in \{1, \dots, M\}^N. \quad (9)$$

The distribution of the received decisions \tilde{U}_j is determined by the conditional probabilities

$$\tilde{\alpha}_{j_k} = P(\tilde{U}_j = k | H_0) = \sum_{l=1}^M T_{kl}^{(j)} \alpha_{j_l}, \quad (10)$$

$$\tilde{\beta}_{j_k} = P(\tilde{U}_j = k | H_1) = \sum_{l=1}^M T_{kl}^{(j)} \beta_{j_l}. \quad (11)$$

Assuming knowledge of the channel transition matrices \mathbf{T}_j , the probability vectors $\tilde{\alpha}_j = \mathbf{T}_j \alpha_j$ and $\tilde{\beta}_j = \mathbf{T}_j \beta_j$ characterizing the distribution of the received local decisions $\tilde{U}_1, \dots, \tilde{U}_N$ under each of the two hypotheses can be calculated.

C. Optimal channel-aware fusion rule

At the fusion center, the received decisions $\tilde{U}_1, \dots, \tilde{U}_N$ are fused to the final detection result $U_0 = \delta_0(\tilde{U}_1, \dots, \tilde{U}_N)$, where the fusion rule δ_0 is a binary-valued mapping

$$\delta_0: \{1, \dots, M\}^N \rightarrow \{0, 1\}. \quad (12)$$

The sensor network detection performance is measured in terms of the global probability of error

$$P_e = \pi_0 P_f + \pi_1 P_m, \quad (13)$$

which can be written as a weighted sum of the global probability of false alarm $P_f = P(U_0 = 1|H_0)$ and the corresponding global probability of miss $P_m = P(U_0 = 0|H_1)$.

The optimal fusion rule under the minimum probability of error criterion can be performed by evaluating a log-likelihood ratio test of the form

$$\sum_{j=1}^N \mathcal{L}_j \begin{matrix} U_0 = 1 \\ \geq \\ U_0 = 0 \end{matrix} \log \left(\frac{\pi_0}{\pi_1} \right) = \vartheta, \quad (14)$$

where $\mathcal{L}_j = \log(P(\tilde{U}_j|H_1)/P(\tilde{U}_j|H_0))$ is the log-likelihood ratio of the received decision \tilde{U}_j and ϑ is the fusion threshold.

It is important to note that once the quantization probabilities (10) and (11) of the received local decisions $\tilde{U}_1, \dots, \tilde{U}_N$ are determined, the optimal channel-aware fusion rule (14) is also determined.

D. Global error probabilities

When using the optimal fusion rule according to (14), the global probability of false alarm P_f and the global probability of miss P_m are determined by the conditional tail probabilities

$$P_f = P\left(\sum_{j=1}^N \mathcal{L}_j > \vartheta | H_0\right) \quad (15)$$

and

$$P_m = P\left(\sum_{j=1}^N \mathcal{L}_j \leq \vartheta | H_1\right). \quad (16)$$

In order to efficiently evaluate the sensor network detection performance in terms of the global probability of error P_e , we employ an approach introduced in [11] which provides tight upper bounds on the global probability of false alarm (15) and the global probability of miss (16).

III. CHERNOFF INFORMATION-BASED OPTIMIZATION OF SENSOR DECISION RULES

In this section, we motivate and present the Chernoff information-based optimization procedure for the local sensor decision rules. The rationale behind this approach is that the Chernoff information arises as asymptotic error exponent in Bayesian hypothesis testing [12].

A. Hypothesis testing and Chernoff information

If we assume conditionally independent and identically distributed (i.i.d.) sensor observations X_1, \dots, X_N , the local conditional probability density functions $f_j(\cdot|H_k)$ are the same for all $j = 1, \dots, N$, and we can write

$$\begin{aligned} H_0: X_j &\sim f_0, \\ H_1: X_j &\sim f_1, \end{aligned} \quad (17)$$

where f_k is here the conditional probability density function under hypothesis H_k for all sensors. The Chernoff information between the two distributions f_0 and f_1 is defined as

$$D^* = C(f_0, f_1) = - \min_{0 \leq t \leq 1} \log \int f_0(x)^t f_1(x)^{1-t} dx. \quad (18)$$

If the fusion center has access to the unquantized and non-distorted observations X_1, \dots, X_N and uses the Bayes optimal decision rule, for the probability of error P_e asymptotically it holds that

$$\lim_{N \rightarrow \infty} \frac{\log P_e}{N} = -D^*. \quad (19)$$

In other words, for large N we obtain

$$P_e \approx \exp(-ND^*), \quad (20)$$

i.e., the Chernoff information D^* is the asymptotic error exponent in minimum probability of error hypothesis testing. Intuitively, in the unquantized case considered above, every sensor contributes with the full Chernoff information D^* to the exponent in (20). The higher the contributed Chernoff information D^* , the lower the global probability of error P_e . This motivates the approach that in the case of M -ary quantization of observations at the sensors, the quantization thresholds at each sensor should be chosen such that the Chernoff information between the probability vectors of quantization probabilities (10) and (11) is maximized.

B. Chernoff information between probability vectors

Analogously to definition (18), the Chernoff information between two probability vectors $\mathbf{p} = (p_1, \dots, p_M)'$ and $\mathbf{q} = (q_1, \dots, q_M)'$ in \mathbb{R}^M is given by

$$D^* = C(\mathbf{p}, \mathbf{q}) = - \min_{0 \leq t \leq 1} \log \sum_{k=1}^M p_k^t q_k^{1-t}. \quad (21)$$

According to their local knowledge, the sensors maximize the Chernoff information $C(\tilde{\alpha}_j, \tilde{\beta}_j)$ between the transformed probability vectors $\tilde{\alpha}_j = \mathbf{T}_j \alpha_j$ and $\tilde{\beta}_j = \mathbf{T}_j \beta_j$. Thereby, we assume that every sensor S_j has knowledge of its own observation statistics given by the conditional marginal probability density functions $f_j(\cdot|H_k)$, $k = 0, 1$, and that it has local channel state information, i.e., it has knowledge of the channel transition matrix \mathbf{T}_j . The observation statistics of the other sensors as well as their channel state information do not have to be available at sensor S_j . Furthermore, the knowledge of the prior probabilities is not necessary at the sensors.

Based on the knowledge available locally at the sensors, the quantization thresholds $\tau_{j1}, \dots, \tau_{jM-1}$ of sensor S_j are

determined in such a way that the Chernoff information between the corresponding probability vectors of quantization probabilities is maximized.

IV. NUMERICAL RESULTS

In the following, we provide a detailed numerical analysis of the Chernoff information-based optimization procedure. First, we consider conditionally i.i.d. observations and compare the detection performance between 1-bit and 2-bit quantization at the local sensors, i.e., we consider distributed detection systems consisting of binary and quaternary sensors, respectively. We study the influence of the local observation signal-to-noise ratio (SNR) on both detection performance and quantization thresholds. Furthermore, the robustness with respect to the underlying prior probabilities is shown. Finally, we consider the case when the observations are non-identically distributed, i.e., we assume that the local observation SNR varies randomly across sensors. For the sake of simplicity, we consider ideal communication channels.

A. Joint distribution of sensor observations

As an illustrative example, we consider the problem of detecting the presence or absence of a deterministic signal in Gaussian noise, i.e., we assume that the random observations X_1, \dots, X_N at the local sensors are distributed according to

$$\begin{aligned} H_0: X_j &\sim \mathcal{N}(0, \sigma_j^2), \\ H_1: X_j &\sim \mathcal{N}(\mu_j, \sigma_j^2), \end{aligned} \quad (22)$$

for $j = 1, \dots, N$. The variance σ_j^2 describes the Gaussian background noise and the mean μ_j indicates the deterministic signal component under hypothesis H_1 at sensor S_j . Accordingly, the local observation SNR at sensor S_j is given by

$$\text{SNR}_j = 10 \log_{10} \left(\frac{\mu_j^2}{\sigma_j^2} \right) \quad [\text{dB}]. \quad (23)$$

The local log-likelihood ratios L_j are again Gaussian random variables with conditional marginal distributions according to

$$\begin{aligned} H_0: L_j &\sim \mathcal{N} \left(-\frac{\mu_j^2}{2\sigma_j^2}, \frac{\mu_j^2}{\sigma_j^2} \right), \\ H_1: L_j &\sim \mathcal{N} \left(\frac{\mu_j^2}{2\sigma_j^2}, \frac{\mu_j^2}{\sigma_j^2} \right). \end{aligned} \quad (24)$$

In the distributed detection systems considered in the following, the local log-likelihood ratios L_j are quantized to 1 or 2 bits, respectively.

B. Optimal centralized detection system

For the optimal centralized detection system that has access to all the unquantized and non-distorted observations X_1, \dots, X_N , the minimum probability of error P_e^* for the detection problem (22) can be calculated explicitly. Calculating the Mahalanobis distance d_M which is given in the present case by

$$d_M = \sqrt{\sum_{j=1}^N \frac{\mu_j^2}{\sigma_j^2}}, \quad (25)$$

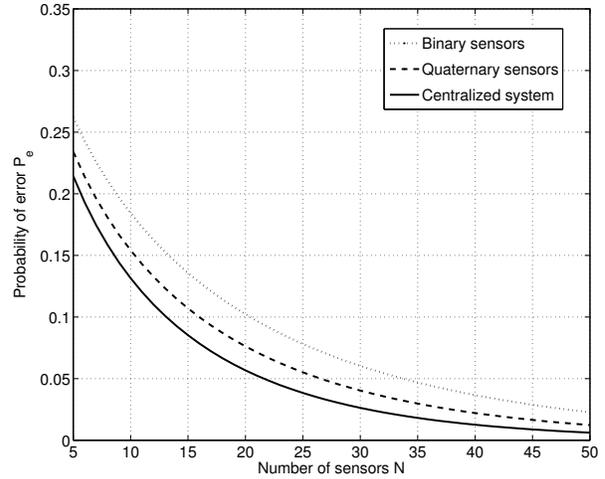


Fig. 2. Probability of error of the Chernoff information-based distributed detection systems with binary and quaternary sensors compared to the optimal centralized detection system at low observation SNR of -3 dB.

the minimum probability of error P_e^* of the optimal centralized detection system can be calculated as

$$P_e^* = \pi_0 \cdot \left(1 - \Phi \left(\frac{\vartheta + \frac{1}{2}d_M^2}{d_M} \right) \right) + \pi_1 \cdot \Phi \left(\frac{\vartheta - \frac{1}{2}d_M^2}{d_M} \right), \quad (26)$$

where Φ is the cumulative distribution function of the standard normal distribution and $\vartheta = \log(\pi_0/\pi_1)$. The minimum probability of error P_e^* of the optimal centralized detection system will be the benchmark for the performance of the Chernoff-information based distributed detection systems. It should be kept in mind however, that the optimal centralized detection system achieving P_e^* has full access to the unquantized and non-distorted observations X_1, \dots, X_N .

C. Conditionally i.i.d. observations

First, we consider the case of conditionally i.i.d. observations and analyze the performance of distributed detection systems with binary sensors and quaternary sensors, respectively. In the case of binary sensors, i.e., $M = 2$, there is only one local quantization threshold τ_{j1} at every sensor. Since the observations are conditionally i.i.d., the optimal threshold maximizing the local Chernoff information is identical for every sensor. In the case of quaternary sensors, i.e., $M = 4$, the same argument holds, so that the three quantization thresholds $\tau_{j1}, \tau_{j2}, \tau_{j3}$ are the same for every sensor. After the thresholds are determined such that the Chernoff information is maximized, we employ a technique presented in [11] in order to numerically evaluate the probability of error P_e of the distributed detection systems with high accuracy.

Fig. 2 illustrates the numerical results at low observation SNR. The probability of error P_e is evaluated for sensor networks consisting of $N = 5, \dots, 50$ sensors for a local observation SNR of -3 dB. The prior probabilities are assumed to be $\pi_0 = \pi_1 = 0.5$. The probability of error of distributed detection systems with quaternary sensors is considerably

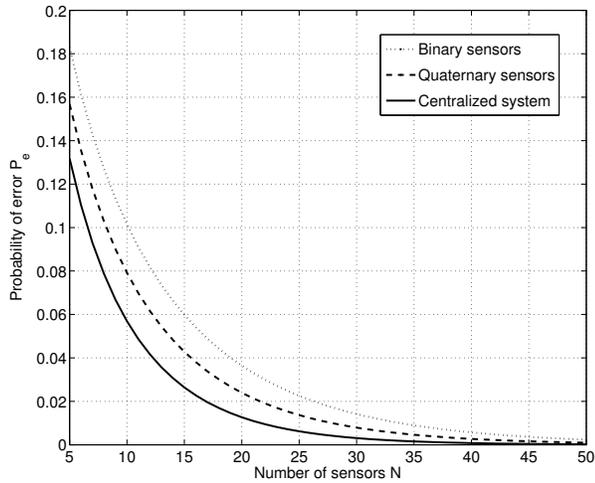


Fig. 3. Probability of error of the Chernoff information-based distributed detection systems with binary and quaternary sensors compared to the optimal centralized detection system at medium observation SNR of 0 dB.

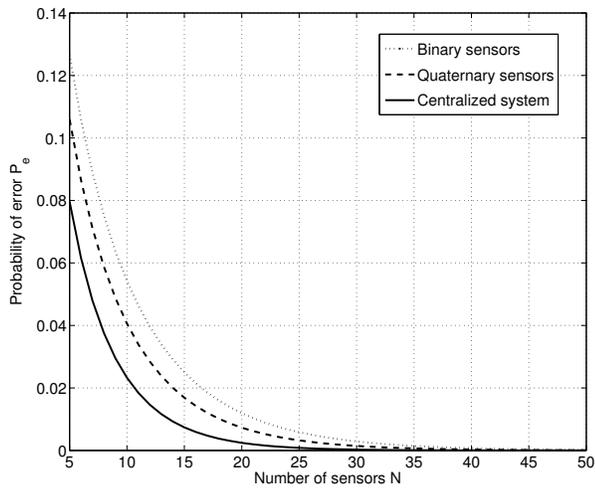


Fig. 4. Probability of error of the Chernoff information-based distributed detection systems with binary and quaternary sensors compared to the optimal centralized detection system at high observation SNR of 2 dB.

smaller compared to distributed detection systems with binary sensors. For example, in order to obtain a global probability of error of $P_e \approx 0.05$, one needs either 22 unquantized observations, 27 quaternary sensors or 33 binary sensors.

Fig. 3 illustrates the numerical results at a medium observation SNR of 0 dB. Again, the probability of error is evaluated for sensor networks consisting of $N = 5, \dots, 50$ sensors and the prior probabilities are assumed to be equal. The probability of error of distributed detection systems with quaternary sensors is approximately in the middle between distributed detection systems with binary sensors and the optimal centralized detection system.

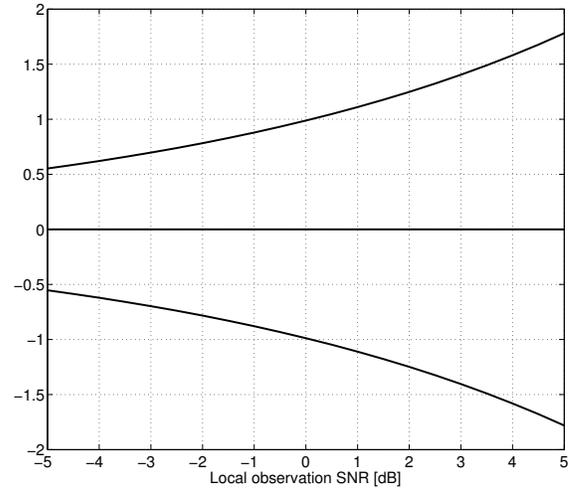


Fig. 5. Optimal quantization thresholds $\tau_{j_1}, \tau_{j_2}, \tau_{j_3}$ of the Chernoff information-based distributed detection system with quaternary sensors as a function of the local observation SNR.

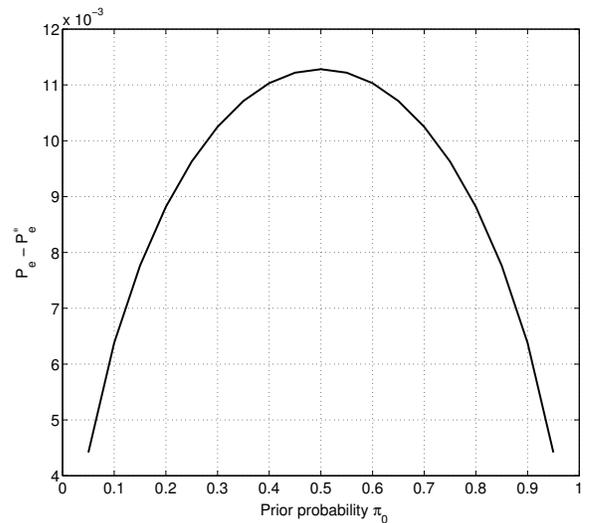


Fig. 6. Difference $P_e - P_e^*$ between the probability of error of the Chernoff-information based distributed detection system with $N = 20$ quaternary sensors and the optimal centralized detection system at a local observation SNR of 0 dB plotted against the prior probability π_0 .

Fig. 4 illustrates the numerical results at a high observation SNR of 2 dB. In order to obtain a global probability of error of $P_e \approx 0.01$, one needs either 14 unquantized observations, 18 quaternary sensors or 21 binary sensors.

The optimal quantization thresholds $\tau_{j_1}, \tau_{j_2}, \tau_{j_3}$ of the Chernoff information-based distributed detection system with quaternary sensors as a function of the local observation SNR are depicted in Fig. 5. The results show that the local log-likelihood ratios L_j are quantized symmetrically, i.e., $\tau_{j_2} = 0$ and $\tau_{j_1} = -\tau_{j_3}$. Furthermore, the higher the local observation SNR, the larger is the distance between the lowest threshold τ_{j_1} and the highest threshold τ_{j_3} . Obviously, the observation

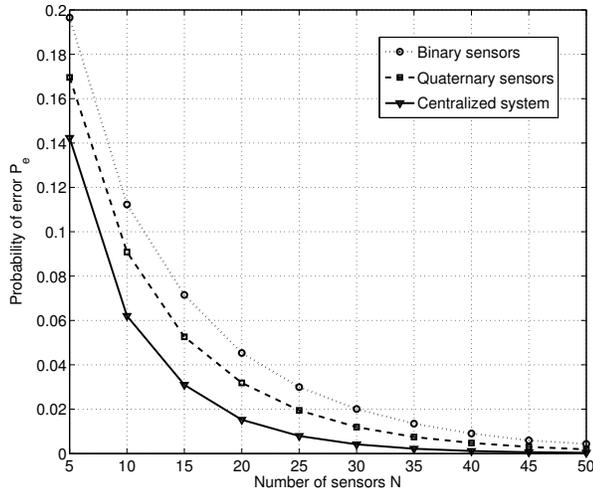


Fig. 7. Probability of error of the Chernoff information-based distributed detection systems with binary and quaternary sensors compared to the optimal centralized detection system for a randomly distributed observation SNR between -3 and 2 dB.

SNR level has a direct influence on local quantizer design.

The robustness of the Chernoff information-based optimization procedure with respect to the prior probabilities is shown in Fig. 6. For a sensor network of $N = 20$ quaternary sensors and a local observation SNR of 0 dB the difference $P_e - P_e^*$ between the probability of error of the Chernoff-information based distributed detection system and the optimal centralized detection system is plotted against the prior probability π_0 . The results show that the maximal deviation between the two probabilities of error is reached for $\pi_0 = \pi_1 = 0.5$, i.e., in our previous numerical results we have already considered the worst case. For unequal prior probabilities, the performance gap between the Chernoff information-based distributed detection systems and the optimal centralized one is even smaller.

D. Non-identically distributed observations

In general, the asymptotic considerations presented in (19) and (20) only hold for conditionally i.i.d. observations. However, we take them as motivation to study the decoupling of the optimization across sensors also for non-identically distributed observations. As in the i.i.d. case, every sensor determines the quantization thresholds in a way that the Chernoff information between the corresponding quantization probabilities is maximized, although the maximum value of the Chernoff information now may vary from sensor to sensor.

Fig. 7 illustrates the numerical results when the local observation SNR is a uniformly distributed random variable between -3 and 2 dB. The probability of error is evaluated for sensor networks consisting of $N = 5, 10, \dots, 50$ sensors and the prior probabilities are assumed to be $\pi_0 = \pi_1 = 0.5$. For the considered scenario, the probability of error of distributed detection systems with quaternary sensors is approximately in the middle between distributed detection systems with binary sensors and the optimal centralized detection system. The

numerical results strongly indicate that Chernoff information-based optimization of sensor networks for distributed detection is also feasible in the case of non-identically distributed observations as long as the observations are conditionally independent.

V. CONCLUSIONS

In this paper, we presented an approach to the optimization of sensor networks for distributed detection that is based on the local maximization of the Chernoff information between the probability vectors of quantization probabilities at every sensor. By considering the problem of detecting a deterministic signal in the presence of Gaussian noise, the numerical results reveal the effect of 1-bit and 2-bit quantization of sensor observations on the overall detection performance when using the Chernoff information-based optimization approach. Furthermore, the dependency of the optimal quantization thresholds with respect to the local observation SNR and the robustness to the prior probabilities are revealed. Finally, it is shown that the Chernoff information-based optimization procedure is also feasible for realistic scenarios with non-identically distributed observations.

ACKNOWLEDGMENT

This work was partly supported by the UMIC excellence cluster of RWTH Aachen University.

REFERENCES

- [1] D. Li, K. D. Wong, Y. H. Hu and A. M. Sayeed, "Detection, classification, and tracking of targets in distributed sensor networks," in *IEEE Signal Processing Magazine*, vol. 19, no. 2, pp. 17-29, March 2002.
- [2] J.-F. Chamberland and V. V. Veeravalli, "Decentralized detection in sensor networks," *IEEE Transactions on Signal Processing*, vol. 51, no. 2, pp. 407-416, Feb. 2003.
- [3] J.-F. Chamberland and V. V. Veeravalli, "Wireless sensors in distributed detection applications," in *IEEE Signal Processing Magazine*, vol. 24, no. 3, pp. 16-25, May 2007.
- [4] B. Chen, L. Tong, and P. K. Varshney, "Channel-aware distributed detection in wireless sensor networks," in *IEEE Signal Processing Magazine*, vol. 23, no. 4, pp. 16-26, July 2006.
- [5] A. R. Reibman and L. W. Nolte, "Optimal detection and performance of distributed sensor systems," in *IEEE Transactions on Aerospace and Electronic Systems*, vol. 23, no. 1, pp. 24-30, January 1987.
- [6] Z.-B. Tang, K. R. Pattipati, and D. L. Kleinman, "An algorithm for determining the decision thresholds in a distributed detection problem," in *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 21, no. 1, pp. 231-237, January/February 1991.
- [7] M. Xiang and C. Han, "Global optimization for distributed and quantized Bayesian detection system," in *Proc. of FUSION'00*, Paris, France, July 2000.
- [8] S. Aldosari and J. M. F. Moura, "Fusion in sensor networks with communication constraints," in *Proc. of IPSN'04*, Berkeley, CA, April 2004.
- [9] Y. I. Han and T. Kim, "Mutual and conditional mutual information for optimizing distributed Bayes detectors," in *IEEE Transactions on Aerospace and Electronic Systems*, vol. 37, no. 1, pp. 147-157, January 2001.
- [10] D. Warren and P. Willett, "Optimum quantization for detector fusion: Some proofs, examples, and pathologies," *J. Franklin Inst.*, vol. 336, pp. 323-359, 1999.
- [11] G. Fabeck and R. Mathar, "Tight performance bounds for distributed detection," in *Proc. of ICASSP'07*, Honolulu, HI, vol. 3, pp. 1049-1052, April 2007.
- [12] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. John Wiley & Sons, 2006.