Power Allocation for Distributed Detection in Wireless Sensor Networks Using Jensen Shannon Divergence

S. Hakimi[1], G. A. Hodtani[2]*

Received :2015\11\25   Accepted:2015\12\25

Abstract

The problem of power allocation for distributed detection in a power constrained wireless sensor network is considered. The sensors are assumed to have independent observations and send their local decisions to a fusion center over multiple access channels. The Jensen-Shannon divergence between the distributions of the detection statistic under different hypotheses is used as a performance criterion. It is demonstrated that by applying the proposed measure power allocation is more efficient comparing to other criteria like mean square error or Jeffry divergence.

Keywords: distance measures, distributed detection, power allocation, wireless sensor networks (WSNs).

I. Introduction

Wireless sensor network (WSN) has become one of the most influential technologies over the last decades [1]. In the distributed sensor network, observation data is collected from different sensors, possibly processed, and transmitted to a fusion center (FC), where the signals received from sensors are combined to produce a final estimate of the observed quantity. The sensors typically have limited energy resources and communication capabilities. Thus, they are required to compress the sensed data and transmit only a short summary. This scenario is called distributed detection [2], the problem of which has been addressed, especially the design of local decision and fusion rules, for a long time [3]-[4].

Most of the previous papers assumed perfect transmission i.e. no error occurs during the transmission. In reality due to power and communication constraints, this assumption may fail. A few researchers studied these imperfect communications between sensors and the FC [5]. Furthermore, most prior works on estimation in distributed sensor networks have focused on the situation, where the parameter(s) of interest are time-invariant, and either deterministic or i.i.d. Gaussian. In [6], the authors studied the linear decentralized estimation of the source vector, where each sensor linearly encoded its observations and the fusion center also applied a linear mapping to estimate the unknown vector signal based on the received messages, with mean squared error (MSE) as the performance criterion. In [7], the best linear unbiased estimator (BLUE) was applied by the FC to generate estimates of the unknown signal. In [8], both the orthogonal and coherent multiple-access channels were considered and two kinds of optimization problems were formulated: MSE minimization under a global sum transmit power constraint, and sum power minimization problem under an MSE constraint. An asymptotic expression for the MSE outage probability was also derived assuming a large number of sensor nodes. In Ref. [2] it was shown when the sensors have independent and identically distributed (i.i.d.) Gaussian or exponential observations and each sensor transmits its data over a MAC with capacity R, having a set of identical binary sensors is asymptotically optimal, as the number of observations per sensor goes to infinity. In [9], the authors studied the decentralized detection and fusion of a Gaussian signal under the assumption of i.i.d. sensor observations and relay-amplifier scheme. Error exponents and resulting bounds were derived for Bayesian fusion performance. In [5], the optimal power allocation strategy was performed using the approximation of Jeffrey (J) divergence between distributions of detection statistics under different hypotheses as the performance measure. In [10], two statistical distance measures, including elemental J-divergence and elemental L2 distance, with closed-form expressions were exploited as performance metrics. However, the noise statistic is considered independent with the diagonal covariance matrix. The channel matrix is assumed to be identity.

In this paper, the Jensen-Shannon (JS) divergence between distributions of detection statistics is evaluated under different hypotheses as a

1. Phd. Student, Department of Electrical Engineering, Ferdowsi University of Mashhad, Mashhad, Iran (hakimi.saeed@stu-mail.um.ac.ir)

2 . Corresponding Author: Associate Professor, Department of Electrical Engineering, Ferdowsi University of Mashhad, Mashhad, Iran (hodtani@um.ac.ir)

performance criterion. This divergence is globally defined, bounded, symmetric, and only vanishes when the two probability distributions are equal. Performing power optimization using the JS-divergence, the transmission gain of sensors is obtained and demonstrated via tables and figures. The rest of the paper is organized as follows. In the next section, the structure of the distributed detection system is described. In Section III, the JS divergence optimization problem is developed. Simulation results are illustrated in Section IV. Finally, the conclusions are presented in section V.

II.System Model

A similar system model as that used in [5] is illustratetd in figure1.

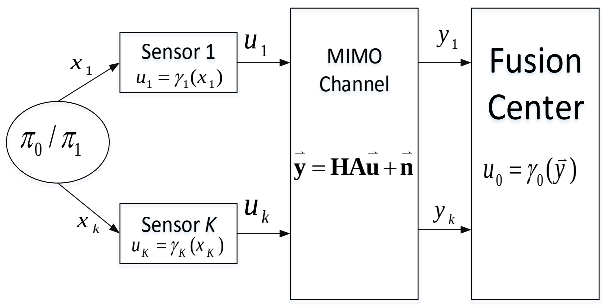

Figure 1. Distributed Detection System Diagram

A binary hypothesis testing problem is considered i.e. at any observation time, the detection of the monitored environment can be summarized as two hypotheses: H_0 and H_1 for the absence and Presence of the target, respectively. Suppose K wireless sensors with independent but not necessarily identically distributed observation $X=[x_1,x_2,...,x_K]$, then based on local binary decision rules as below

$$u_k=\gamma_k(x_k)=\begin{cases}0 & \text{under } H_0\\ 1 & \text{under } H_1\end{cases} \quad (1)$$

sensors make decisions $\mathrm{u}=[u_1,...,u_K]$,and send them to FC. In this work, multi-access channels are considered for communication between the sensors and the FC. The false alarm probability and detection probability of sensor k are given by

$$P_F(k)=p(u_k=1|H_0) \quad (2)$$

and

$$P_D(k)=p(u_k=1|H_1) \quad (3)$$

Accordingly, the joint conditional density functions of the local decisions are

$$p(\mathrm{u}|H_0)=\prod_{k=1}^{K}P_F(k)^{u_k}(1-P_F(k))^{(1-u_k)} \quad (4)$$

and

$$p(\mathrm{u}|H_1)=\prod_{k=1}^{K}P_D(k)^{u_k}(1-P_D(k))^{(1-u_k)} \quad (5)$$

The received signals at FC are

$$\mathrm{y}=\mathbf{HA}\mathrm{u}+\mathrm{n} \quad (6)$$

where $\mathbf{y}=[y_1,...,y_K]$ is the received signal, $\mathbf{H}$ is the channel matrix , which is assumed to be deterministic in this paper (We focus on the case in which the sensors and the FC have the minimal movement and the environment changes slowly). $\mathbf{A}=\mathrm{diag}\{a_1,...,a_K\}$ is the transmission gain matrix for each sensor, and $\mathrm{n}=\{n_1,...,n_K\}$ is the additive Gaussian noise vector with zero mean and covariance matrix R . Thus, the conditional density function of the received signals at FC given the transmitted signals from sensors is multivariate Gaussian as follows

$$p(y|u)=\frac{1}{|2\pi R|^{\frac{1}{2}}}\exp\left[-\frac{1}{2}(y-\mathrm{HAu})^T R^{-1}(y-\mathrm{HAu})\right] \quad (7)$$

The conditional probability density function of received signals given hypotheses is

$$p(y|H_i)=\sum_u p(y|u)p(u|H_i),\ \ i=0,1. \quad (8)$$

The decision rule $u_0=\gamma_0(y)$ at FC determines the final decision. In this paper, the analysis is based on the power control strategy i.e. assigning a partial of the total power budget P_{tot} to each sensor by choosing a transmission gain matrix A

in the presence of total and individual power constraints.

III. Power Allocation Based On JS Divergenece

In this section, a detection performance criterion for power allocation problem is proposed. Error probability is the most commonly used measure of detection performance. However, only in certain special cases it provides the exact probability of error for the optimal detector. Even if the problem reaches the closed-form, usually numerical computations are required to provide insights into performance trends. Consequently, it does not lead to an efficient metric for the considered system. Another method commonly used in detection performance criterion is error bounds [11]. Distance related bounds are among the commonly used detection performance metrics [5]. The J-divergence, first proposed by Jeffrey [12], is a widely used measure for detection performance [11, 13, 14]. It symmetrizes the oriented Kullback-Leibler (KL) divergence as follows:

$$\begin{aligned} J(p_1,p_0) &= J(p_0,p_1) = KL(p_1,p_0)+KL(p_0,p_1) \\ &= H(p_1,p_0)+H(p_0,p_1)-(H(p_1)+H(p_0)) \end{aligned} \quad (9)$$

where $KL(p_1,p_0)$ is the KL divergence between p_1 and p_0

$$KL(p_1,p_0)=\int p_1 \log\frac{p_1(x)}{p_0(x)}dx \quad (10)$$

The J-divergence puts a lower bound on the detection error probability [11] via the following inequality

$$P_e \geq \pi_1\pi_2 \exp(-J/2) \quad (11)$$

where $\pi=(\pi_1,\pi_2)$ is the prior probability set, and π_i is the prior probability of the hypothesis H_i .One drawback of Jeffrey's and KL divergence is that it may be unbounded, as the Pinsker inequality stated [15], and as a result numerically quite unstable to compute in practice. In this paper, the JS divergence is applied as the performance criterion. The JS-divergence was introduced in [16]:

$$JS(p_0,p_1)=H(\pi_1 p_1+\pi_2 p_2)-\pi_1 H(p_1)-\pi_2 H(p_2) \quad (12)$$

For the special case which $\pi_1=\pi_2=\frac{1}{2}$, we have

$$\begin{aligned} &JS_{1/2}(p_1,p_0)=JS_{1/2}(p_0,p_1)= \\ &\frac{1}{2}\left(KL(p_1,\frac{p_1+p_0}{2})+KL(p_0,\frac{p_1+p_0}{2})\right) \end{aligned} \quad (13)$$

It symmetrizes the KL divergence by taking the average relative entropy of each distribution to the entropy of the average distribution $\frac{p_1+p_0}{2}$.

JS-divergence is bounded [16] and its square root which obeys the triangular inequality serves as a metric [17]. By the inequality of the arithmetic and geometric means, we have

$$\frac{p_1+p_0}{2} \geq \sqrt{p_1 p_0} \quad (14)$$

thus

$$KL(p_1,\frac{p_1+p_0}{2})=\int p_1\log\frac{p_1}{\frac{p_1+p_0}{2}} \leq \int p_1 \log\frac{p_1}{\sqrt{p_1p_0}}=\frac{1}{2}KL(p_1,p_0) \quad (15)$$

consequently

$$JS_{1/2}(p_1,p_0) \leq \frac{1}{4}J(p_1,p_0) \quad (16)$$

Therewith, the JS-divergence is a lower bound on J-divergence. It has been shown that the JS-divergence provides both the lower and upper bounds for the Bayes probability of error [16].

$$\begin{aligned} &\frac{1}{4}(H(\pi_0,\pi_1)-JS_{1/2}(p_0,p_1))^2 \\ &\leq p_e(p_0,p_1) \leq \\ &\frac{1}{2}(H(\pi_0,\pi_1)-JS_{1/2}(p_0,p_1)) \end{aligned} \quad (17)$$

where H is the Shannon entropy function and $H(\pi_0,\pi_1)=-\pi_0\log\pi_0-\pi_1\log\pi_1$ for which $\pi_0=\pi_1=\frac{1}{2}$ (17)results in the following inequality

$$\frac{1}{4}(1-JS_{1/2}(p_0,p_1))^2 \leq p_e(p_0,p_1) \leq \frac{1}{2}(1-JS_{1/2}(p_0,p_1)) \quad (18)$$

No Similar upper bound on p_e in terms of the J-divergence appears to be generally true. Thus, this measure is particularly useful in the study of decision problems.

In this work, to optimize the detection performance at FC, the JS-divergence computed for the two densities of the received signals is maximized with respect to the underlying hypotheses. Thus, solving the following optimization problem results in the optimal power allocation:

$$\max JS_{1/2}\left(p(y\mid H_1), p(y\mid H_0)\right)$$
$$\text{s.t } \mathrm{Tr}\left[\mathbf{A}\mathbf{A}^T\right] \le P_{tot} \qquad (18)$$
$$0 \le \mathbf{A} \le \sqrt{P_{\max}}$$

where P_{tot} is the total transmitter power budget distributed among sensors, and $\sqrt{P_{\max}}$ is the component-wise square root of individual power constraints. The JS divergence $JS = JS_{1/2}\left(p(y\mid H_1), p(y\mid H_0)\right)$ is given by

$$JS\left(p(y\mid H_1), p(y\mid H_0)\right) = \int_y dy \left[p(y\mid H_1)\log\frac{p(y\mid H_1)}{\frac{1}{2}p(y\mid H_1)+\frac{1}{2}p(y\mid H_0)} + p(y\mid H_0)\log\frac{p(y\mid H_0)}{\frac{1}{2}p(y\mid H_1)+\frac{1}{2}p(y\mid H_0)} \right] \qquad (20)$$

The density functions $p(y\mid H_i),\ i\in\{0,1\}$ are given by (8). These conditional density functions $p(y\mid H_i)$ are Gaussian mixtures. Generally, most well known divergences including KL divergence do not yield an analytic closed-form expression for a mixture of Gaussians. Similarly, the JS-divergence between two Gaussian mixture densities does not have a general closed-form expression, so some approximations are needed to present the objective function in the closed-form. Different techniques have been used in literature to work around this problem [18]. Some approaches estimate the divergences. In this paper, a commonly used approximation that replaces the Gaussian mixtures densities $p(y\mid H_i)$ with Gaussian densities $p_g(y\mid H_i)=\mathcal{N}(y;\mu_i,\Sigma_i)$ is applied. In this technique, f and g are considered Gaussian mixture models, and marginal densities of $x\in R^d$ under f and g are $f(x)=\sum_\alpha \omega_\alpha \mathcal{N}(x;\mu_\alpha,\Sigma_\alpha)$ and $g(x)=\sum_b \psi_b \mathcal{N}(x;\mu_b,\Sigma_b)$, where ω_α is the prior probability of each state, and $\mathcal{N}(x;\mu_\alpha,\Sigma_\alpha)$ is a Gaussian in x with mean μ_α and variance Σ_α.Replacing f and g with Gaussians $\hat{f}$ and $\hat{g}$ whose moments matches, the mean and covariance of f are given by

$$\mu_{\hat{f}} = \sum_\alpha \omega_\alpha \mu_\alpha \qquad (19)$$

$$\Sigma_{\hat{f}} = \sum_\alpha \omega_\alpha (\Sigma_\alpha + (\mu_\alpha-\mu_{\hat{f}})(\mu_\alpha-\mu_{\hat{f}})^T) \qquad (20)$$

From (7) and (8), it is clear that $p(\mathbf{u}\mid H_i)$ is the prior probability of each state, and $p(y\mid\mathbf{u})$ is a Gaussian in y with mean $\mathbf{HAu}$ and covariance matrix $\mathbf{R}$. Therefore, using the mentioned technique, the mean and covariance of Gaussian densities are given by

$$\mu_i = \sum_u \mathbf{HAu}p(\mathbf{u}\mid H_i) \qquad (21)$$

$$\Sigma_i = \sum_u \left(\mathbf{R} + (\mathbf{HAu}-\mu_i)(\mathbf{HAu}-\mu_i)^T\right) p(\mathbf{u}\mid H_i) \qquad (22)$$

Applying (4) and (5), we have

$$\mu_i = \mathbf{HAu}\beta_i \ \text{ for } i\in\{0,1\} \qquad (23)$$

$$\beta_1 = \sum_{\mathbf{u}} \mathbf{u}p(u\mid H_1) = \left[P_D(1),\dots,P_D(K)\right]^T \qquad (24)$$

$$\beta_0 = \sum_{\mathbf{u}} \mathbf{u}p(u\mid H_0) = \left[P_F(1),\dots,P_F(K)\right]^T \qquad (25)$$

$$\Sigma_i = \mathbf{R} + \mathbf{HAB}_i\mathbf{A}^T\mathbf{H}^T \ \text{ for } i\in\{0,1\} \qquad (26)$$

$$\mathbf{B}_1 = diag\left\{P_D(1)\left[1-P_D(1)\right],\dots,P_D(K)\left[1-P_D(K)\right]\right\} \qquad (27)$$

$$\mathbf{B}_0 = diag\left\{P_F(1)\left[1-P_F(1)\right],\dots,P_F(K)\left[1-P_F(K)\right]\right\} \qquad (28)$$

Unfortunately, the JS divergence has no closed-form for Gaussians because $KL(p_{g_1},\frac{p_{g_1}+p_{g_0}}{2})$ and $KL(p_{g_0},\frac{p_{g_1}+p_{g_0}}{2})$ do not have closed form expressions for Gaussians. It is due to that $\frac{p_{g_1}+p_{g_0}}{2}$ is the mixture distribution of two Gaussians. Again, by applying the mentioned technique the $p_{g_m}=\frac{p_{g_1}+p_{g_0}}{2}$ part is approximated by Gaussians whose mean and covariance are given by

$$\mu_m = \frac{\mu_1+\mu_2}{2} \qquad (29)$$

and

$$\Sigma_m = \frac{1}{2}\left(\Sigma_0+\Sigma_1+\frac{1}{2}(\mu_0-\mu_1)(\mu_0-\mu_1)^T\right) \qquad (30)$$

Next, the JS-divergence between Gaussian densities is derived

$$JS(p_{g_1}, p_{g_0}) = \frac{1}{2}\left(KL(p_{g_1}, p_{g_m}) + KL(p_{g_0}, p_{g_m})\right) \quad (31)$$
$$= \frac{1}{4}\log\det \Sigma_m \Sigma_1^{-1}$$
$$+\frac{1}{4}\mathrm{Tr}\Sigma_m^{-1}\left[(\mu_1-\mu_m)(\mu_1-\mu_m)^T + \Sigma_1 - \Sigma_m\right]$$
$$+\frac{1}{4}\log\det \Sigma_m \Sigma_0^{-1}$$
$$+\frac{1}{4}\mathrm{Tr}\Sigma_m^{-1}\left[(\mu_0-\mu_m)(\mu_0-\mu_m)^T + \Sigma_0 - \Sigma_m\right]$$

Applying the (29) and (30), we have

$$JS(p_{g_1}, p_{g_0}) = \frac{1}{4}\log\det \Sigma_m \Sigma_1^{-1}\Sigma_0^{-1}\Sigma_m{}^T \quad (32)$$
$$+\frac{1}{4}\mathrm{Tr}\Sigma_m^{-1}\left(\Sigma_0 + \Sigma_1 - 2\Sigma_m + \frac{1}{2}(\mu_0-\mu_1)(\mu_0-\mu_1)^T\right)$$
$$= \frac{1}{4}\log\det \Sigma_m \Sigma_1^{-1}\Sigma_0^{-1}\Sigma_m{}^T$$
$$= \frac{1}{4}\log\det \frac{1}{4}\left[\Sigma_0 + \Sigma_1 + \frac{1}{2}\mu\mu^T\right]\Sigma_1^{-1}\Sigma_0^{-1}\left[\Sigma_0 + \Sigma_1 + \frac{1}{2}\mu\mu^T\right]^T$$

where $\mu = \mu_1 - \mu_0$.

Therefore, a power allocation strategy can be designed by solving the following optimization problem

$$\max JS(p_{g_1}, p_{g_0})$$
$$s\,t \quad \mathrm{Tr}[\mathbf{A}\mathbf{A}^T] \le P_{tot}$$
$$0 \le \mathbf{A} \le \sqrt{P_{\max}}$$

where the optimization is over amplitude matrix $\mathbf{A}$ defined earlier.

IV. Simulation

This section presents simulation results obtained from the considered power allocation strategy. In simulations, all the sensors and the FC perform Neyman-Pearson detection with false alarm probabilities set to $P_F = 0.04$. Both the multiple access and orthogonal channel are considered in different scenarios. The maximum transmitting power of each sensor is $P_{\max} = 2$ mW (3 dBm). The total power budget is below or equal to 2K mW. The detection probabilities vary according to the quality of the received local signal. The interior point optimization algorithm is used to solve the proposed power allocation for all cases. To illustrate channel power gains for each sensor $g_k = -PL_k$, the path loss of signal power (measured in decibels) is modeled as follows:

$$PL_k = PL_0 + 10n\log_{10}(d_k / d_0) \quad (33)$$

where n is path loss exponent set to 2 for free space propagation, PL_0 is a constant set to $55dB$, and d_0 is also a constant set to 1m in simulations.

A. Two Sensors Orthogonal Channels

Two sensors located at $d_1 = 2$ m and $d_2 = 5$ m away from FC are considered to monitor an event and independently make their local decisions, then transmit their local decisions to FC through orthogonal channels. Based on mentioned assumptions, the channel gains are -61 and -69 dB, respectively. The noise covariance matrix at FC is $\mathbf{R} = \sigma^2\mathbf{I}_k$ where the noise variance is $\sigma^2 = -70$ dBm. In this scenario, the total power budget varies from -14 to 6 dBm. Four cases with various local detection quality combinations will be considered.

Case A1: $P_D(1) = 0.9,\ P_D(2) = 0.7$

Case A2: $P_D(1) = 0.9,\ P_D(2) = 0.9$

Case A3: $P_D(1) = 0.1,\ P_D(2) = 0.9$

Case A4: $P_D(1) = 0.7,\ P_D(2) = 0.9$

For comparison, an equal power allocation is also considered. The optimal power allocation as well as the equal power allocation for cases A1-A4 are shown in 0For case A1, when the total power budget is very low, all the power goes to sensor 1which has a better channel and also better detection quality. Growing the power budget, when the sensor 1 reaches maximum output level, sensor 2 starts to get positive power allocation. For case A2, sensor 1 has a better communication channel and the same local detection quality as sensor 2, so like the previous case power allocates to sensor 1 until it reaches maximum output level, after that sensor 2 gets positive power. For case A3, although sensor 1 is closer to FC and has a better channel, the detection quality of sensor 2 is much better, therefore in low power budget no power is distributed to sensor 1until sensor 2 reaches the maximum output power. For case A4, the sensor 2 does not have a noticeable better detection quality, while the better communication channel of sensor 1 predominates. The proposed allocation distributes all the power to sensor 1until sensor 1 reaches the maximum output power and then sensor 2 starts to get power allocation. The water-filling effect of the proposed power allocation is obvious in this scenario. The equal power allocation does not change between the four cases because it is not influenced by the local detection quality.

Figures 3-6 show the detection probability at FC as a function of the total power budget for the cases A1-A4. It can be seen that for the same detection probability the proposed power allocation can save 3 dB in total power compared to equal power allocation.

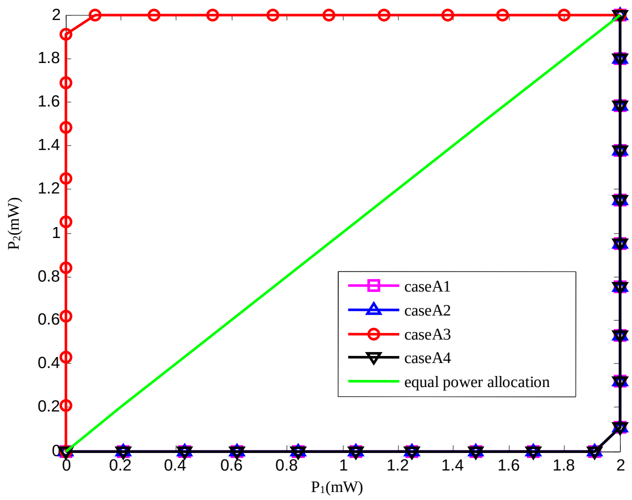

Figure 2. Equal power allocation and the proposed power allocation for the cases A1-A4.

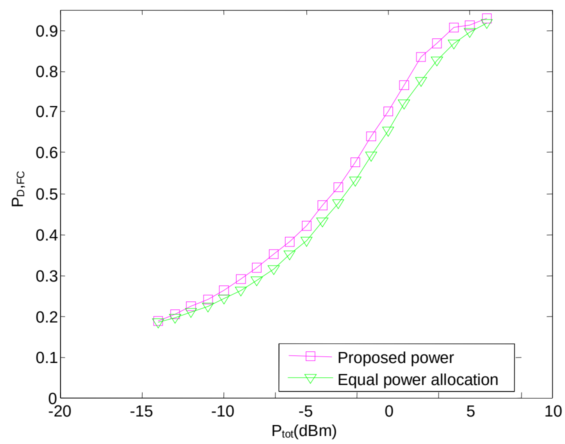

Figure 3. FC detection probability as a function of total power budget for caseA1.

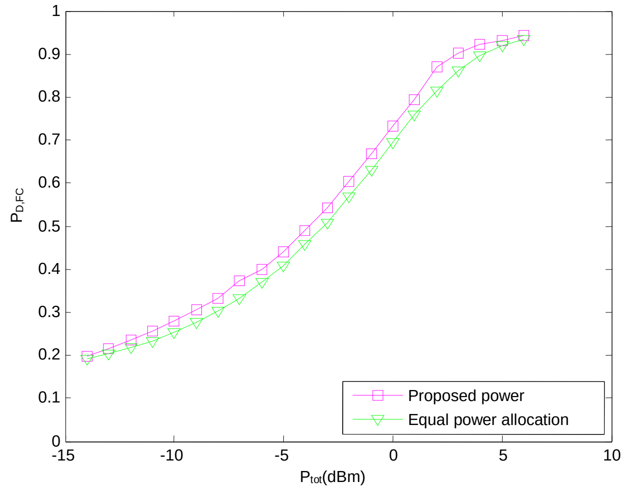

Figure 4. FC detection probability as a function of total power budget for caseA2.

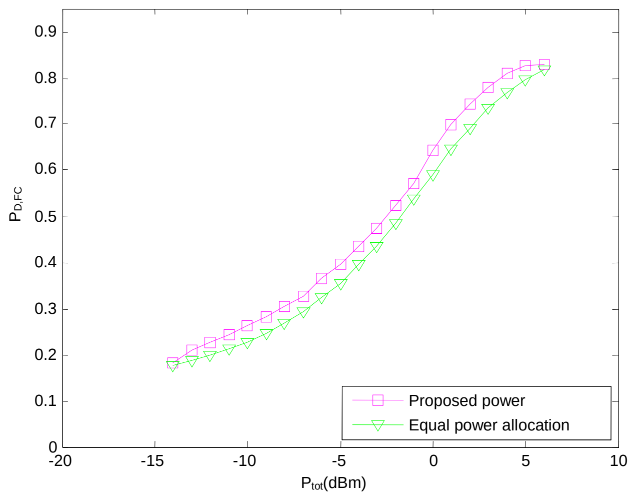

Figure 5. FC detection probability as a function of total power budget for caseA3.

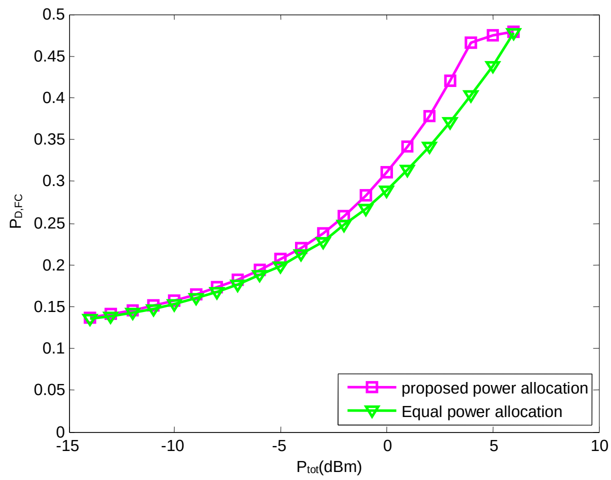

Figure 6. FC detection probability as a function of total power budget for caseA4.

B. Ten sensors with orthogonal channels

In this part, we consider ten sensors scattered around an FC, and the total power budget varies from -7 dBm to 13 dBm (when each sensor transmits at full power 2 mW). Four cases are examined with numerous detection qualities located at various distances.

Case B1: $d_i = 2 + 0.6(i-1),\ P_D(i) = 0.1 + 0.09(i-1)$

Case B2: $d_i = 2 + 0.6(i-1),\ P_D(i) = 0.4 + 0.06(i-1)$

CaseB3: $d_i = 2 + 0.6(i-1),\ P_D(i) = 0.8$

Case B4:

$d_i = 2 + 0.6(i-1),\ P_D(i) = 0.94 - 0.06(i-1)$

Tables I- IV show the percentage of the total power allocated to each sensor for cases B1-B4.

Table 1. Percentage of The Tptal Power Allocated To Each Sensor for Case Bl

P_{tot} (dBm)	1	2	3	4	5	6	7	8	9	10
-7 m	0	0	0	0	0	0	0	0	0	100
-2	0	0	0	0	0	0	0	0	25	75
3	0	0	0	0	0	0	0	15	31	54
8	0	0	0	0	0	7.5	13.5	20	28	31
13	10	10	10	10	10	10	10	10	10	10

Table2. Percentage of The Total Power Allocated to Each Sensor for Case B2

P_{tot} (dBm)	1	2	3	4	5	6	7	8	9	10
-7	100	0	0	0	0	0	0	0	0	0
-2	75	25	0	0	0	0	0	0	0	0
3	41	31	19	9	0	0	0	0	0	0
8	19	17	15	12	10	9	7	5	4	2
13	10	10	10	10	10	10	10	10	10	10

Table3. Percentage of The Total Power Allocated to Each Sensor for Case B3

P_{tot} (dBm)	1	2	3	4	5	6	7	8	9	10
-7	100	0	0	0	0	0	0	0	0	0
-2	67	33	0	0	0	0	0	0	0	0
3	42	34	22	2	0	0	0	0	0	0
8	24	24	21	17	11	3	0	0	0	0
13	10	10	10	10	10	10	10	10	10	10

Table 4. Percentage of the Total Power Allocated to Each Sensor for Case B3

P_{tot} (dBm)	1	2	3	4	5	6	7	8	9	10
-7	100	0	0	0	0	0	0	0	0	0
-2	79	21	0	0	0	0	0	0	0	0
3	53	35	12	2	0	0	0	0	0	0
8	31	30	23	14	2	0	0	0	0	0
13	10	10	10	10	10	10	10	10	10	10

In Case B1, sensors farther from FC have better detection qualities. It can be seen that at low total power budget power initially goes to farther sensors. As the power budget grows, power starts to be allocated to closer sensors and be distributed more evenly among the sensors. When the total power budget reaches the maximum (13 dBm), each sensor transmits at $P_{\max} = 3$ dBm. For case B2, even though sensors farther from FC have better detection qualities, this advantage cannot outperform the better communication channel of closer sensors to FC. Table II shows for low total power budget closer sensors to FC gain more power. When these sensors reach maximum output power, farther sensors get positive power allocation in sequential fashion. For case B3, all sensors have the same detection quality, but the sensors closer to FC have the advantage of better channel gain, so it is not surprising that the closer sensors to FC get positive power at first. In case B4, closer sensors have better detection quality and also better communication channel. This case is a more extreme version of case B3 in term of detection probability. In low power budget, no power is distributed to farther sensors. Power starts to be distributed to farther sensors in high power budget.

C. Two sensors with multi-access channels

The description and quality of the sensors of this part is similar to the section IV-A, but the data transmission is over a multi-access channel. The channel matrix is given by

$$\mathrm{H} = \begin{bmatrix} g_1 & 0 \\ 0 & g_2 \end{bmatrix} \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$$

where g_1 and g_2 are the channel gains and the same as the section IV-A. $\rho = 0.1$ is the interference coefficient. Furthermore, we consider four cases with the local detection combinations exactly the same as section IV-A.

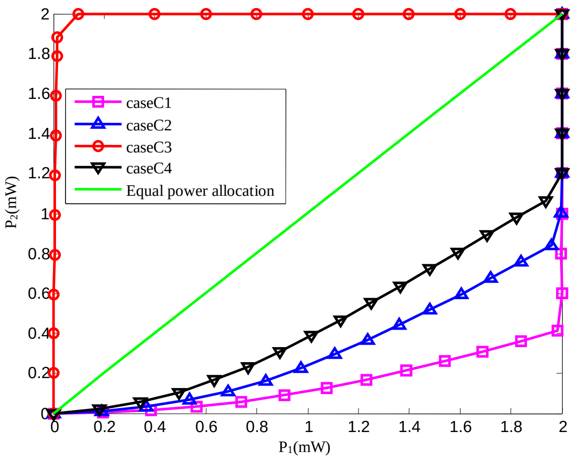

Figure 7. Equal power allocation and the proposed power allocation for the cases C1-C4.

Fig. 7 shows the proposed power allocation as well as the equal power allocation. It is noticed that the optimal power allocation is different from that of orthogonal channels. The water-filling effect in this section is not as obvious as in section IV-A. The interference of the multi-access channel affects the quality of communication channels. It makes the contribution of each sensor at FC depend not only on itself but also on other sensors.

V. Conclusion

A power allocation scheme for distributed detection systems over non ideal communication channels was introduced. The proposed power allocation strategy was analyzed with the JS divergence as the detection performance criterion, which reached the closed form expression. The optimal power allocation by maximizing the proposed criteria depends on channel quality and local detection probability. This criterion outperforms previously used divergences in term of boundedness, the feature that is a valuable property when numerical applications are considered. Furthermore, as demonstrated the JS-divergence is an efficient tool to study decision problems.

References

[1] Q. Z. A. Swami, and Y. Hong, Wireless Sensor Networks: Signal Processing and Communications: *Wiley*, 2007.

[2] J. F. Chamberland and V. V. Veeravalli, "Decentralized detection in sensor networks," *IEEE Transactions on Signal Processing*, vol. 51, pp. 407-416, 2003.

[3] R. Viswanathan and P. K. Varshney, "Distributed detection with multiple sensors I. Fundamentals," *Proceedings of the IEEE*, vol. 85, pp. 54-63, 1997.

[4] R. R. Tenney and N. R. Sandell, "Detection with Distributed Sensors," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-17, pp. 501-510, 1981.

[5] Z. Xin, H. V. Poor, and C. Mung, "Optimal Power Allocation for Distributed Detection Over MIMO Channels in Wireless Sensor Networks," *IEEE Transactions on Signal Processing*, vol. 56, pp. 4124-4140, 2008.

[6] X. Jin-Jun, C. Shuguang, L. Zhi-Quan, and A. J. Goldsmith, "Linear Coherent Decentralized Estimation," *IEEE Transactions on Signal Processing*, vol. 56, pp. 757-770, 2008.

[7] C. Shuguang, X. Jin-Jun, A. J. Goldsmith, L. Zhi-Quan, and H. V. Poor, "Estimation Diversity and Energy Efficiency in Distributed Sensing," *IEEE Transactions on Signal Processing*, vol. 55, pp. 4683-4695, 2007.

[8] A. S. Leong, S. Dey, and J. S. Evans, "Asymptotics and Power Allocation for State Estimation Over Fading Channels," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 47, pp. 611-633, 2011.

[9] S. K. Jayaweera, "Bayesian Fusion Performance and System Optimization for Distributed Stochastic Gaussian Signal Detection Under Communication Constraints," *IEEE Transactions on Signal Processing*, vol. 55, pp. 1238-1250, 2007.

[10] K. Rajan and B. Natarajan, "A Distance Based Comparison of Optimal Power Allocation to Distributed Sensors in a Wireless Sensor Network," in ICC '08. *IEEE International Conference on Communications*, 2008, pp. 4391-4395.

[11] H. Kobayashi and J. B. Thomas, "Distance measures and related criteria," *in 5th Annu. Allerton Conf. Circuit System Theory*, 1967, pp. 491-500.

[12] H. Jeffreys, "An invariant form for the prior probability in estimation problems," *in Roy. Soc. A.*, 1946, pp. 453-461.

[13] T. Kailath, "The Divergence and Bhattacharyya Distance Measures in Signal Selection," *IEEE Transactions on Communication Technology*, vol. 15, pp. 52-60, 1967.

[14] H. Kobayashi, "Distance measures and asymptotic relative efficiency," *IEEE Transactions on Information Theory*, vol. 16, pp. 288-291, 1970.

[15] I. Sason, "On Reverse Pinsker Inequalities," *ArXiv*, May 2015.

[16] L. Jianhua, "Divergence measures based on the Shannon entropy," *IEEE Transactions on Information Theory*, vol. 37, pp. 145-151, 1991.

[17] D. M. Endres and J. E. Schindelin, "A new metric for probability distributions," *IEEE Transactions on Information Theory*, vol. 49, pp. 1858-1860, 2003.

[18] J. R. Hershey and P. A. Olsen, "Approximating the Kullback Leibler Divergence Between Gaussian Mixture Models," *in Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on*, 2007, pp. IV-317-IV-320.