EH-Enabled Distributed Detection Over Temporally Correlated Markovian MIMO Channels

Ghazaleh Ardeshiri, Azadeh Vosoughi *Senior Member, IEEE*
Email:gh.ardeshiri@knights.ucf.edu, azadeh@ucf.edu

ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) | 978-1-7281-6327-7/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICASSP49357.2023.10096256

***Abstract*—We address distributed detection problem in a wireless sensor network, where each sensor harvests and stores randomly arriving energy units in a finite-size battery. Sensors transmit their symbols simultaneously to a fusion center (FC) with M >1 antennas, over temporally correlated fading channels. To characterize the channel time variation we adopt a Markovian model and assume that the channel time-correlation is defined by Jakes-Clark's correlation function. We consider limited feedback of channel gain, defined as the Frobenius norm of MIMO channel matrix, at a fixed feedback frequency (e.g., every T time slots) from the FC to sensors. Modeling the randomly arriving energy units as a Poisson process and the quantized channel gain and the battery dynamics as homogeneous finite-state Markov chains, we propose an adaptive transmit power control strategy such that the J-divergence based detection metric is maximized at the FC, subject to an average transmit power per-sensor constraint.**

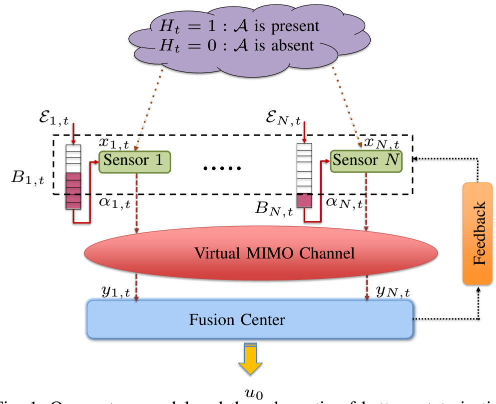

Fig. 1: Our system model and the schematic of battery state in time slot t.

I. Introduction

Event detection is one of the vital tasks in wireless sensor networks (WSNs). Providing a guaranteed detection performance by a conventional WSN, in which sensors are powered by non-rechargeable batteries and become inactive when their stored energy is exhausted, is unfeasible [1]–[13]. Energy harvesting (EH) from the environment (e.g., ambient RF and renewable energy sources) is a promising solution to address the energy constraint problem in conventional WSNs [15]–[19]. In EH-powered WSNs power/energy management is necessary, in order to balance the rates of energy harvesting and energy consumption for transmission. Since ambient RF and renewable energy sources are intrinsically time-variant and sporadic, stochastic models are suitable to model randomly arriving energy and harvested energy. In addition, wireless communication channels change randomly in time due to fading. These together prompt the need for developing new adaptive transmit power control strategies for an EH-enabled transmitter that can adapt to the random energy arrivals and time-varying fading channels (according to the limited channel state information (CSI) available through feedback channel) such that a certain detection performance is guaranteed.

Considering a WSN, composed of EH-enabled sensors and a fusion center (FC), in [16] we developed an *adaptive channel-dependent transmit power control strategy* for sensors such that J-divergence detection metric at the FC is maximized. In [16] we assumed that the FC has a single antenna, sensors transmit over orthogonal channels, and fading channels between sensors and the FC are independent and identically distributed (i.i.d.) over time slots. In this work we extend [16] to *temporally correlated Markovian MIMO channels* where the FC has $M > 1$ antennas, sensors transmit their symbols to the FC *simultaneously*, and fading channels between sensors and the FC are *correlated over time slots*. Our proposed power control strategy allows each sensor to adapt its transmit power in each time slot, based on its current battery state and the latest available channel gain feedback.

II. System Model

A. Observation Model at Sensors

Suppose the time horizon is divided into slots of equal length T_s. Each time slot is indexed by an integer t for $t = 1, 2, ..., \infty$. We model the underlying binary hypothesis H_t in time slot t as a binary random variable $H_t \in \{0, 1\}$ with a-priori probabilities $\Pi_0 = \Pr(H_t = 0)$ and $\Pi_1 = \Pr(H_t = 1) = 1 - \Pi_0$. We assume that the hypothesis H_t varies over time slots in an independent and identically distributed (i.i.d.) manner. Let $x_{n,t}$ denote the local observation at sensor n in time slot t. We assume that sensors' observations given each hypothesis with conditional distribution $f(x_{n,t}|H_t = h_t)$ for $h_t \in \{0, 1\}$ are independent across sensors. This model is relevant for WSNs that are tasked with detection of a known signal in uncorrelated Gaussian noises with the following signal model

$$\begin{aligned} H_t = 1: &\quad x_{n,t} = \mathcal{A} + v_{n,t}, \\ H_t = 0: &\quad x_{n,t} = v_{n,t}, \quad \text{for } n = 1, \dots, N, \end{aligned} \tag{1}$$

where Gaussian observation noises $v_{n,t} \sim \mathcal{N}(0, \sigma^2_{v_n})$ are independent over time slots and across sensors. Given observation $x_{n,t}$ sensor n computes its local log-likelihood ratio (LLR)

$$\xi_n(x_{n,t}) \triangleq \log\left(\frac{f(x_{n,t}|H_t = 1)}{f(x_{n,t}|H_t = 0)}\right), \tag{2}$$

and compares it against a given local threshold θ_n to choose its non-negative transmission symbol $\alpha_{n,t}$. When $\xi_n(x_{n,t}) < \theta_n$, sensor n lets $\alpha_{n,t} = 0$. When $\xi_n(x_{n,t}) > \theta_n$, sensor n chooses $\alpha_{n,t}$ according to the rule in (11). We have

$$\widehat{\Pi}_{n,0} = \Pr(\alpha_{n,t}=0) = \Pi_0(1-P_{\mathrm{f}_n}) + \Pi_1(1-P_{\mathrm{d}_n}),$$
$$\widehat{\Pi}_{n,1} = \Pr(\alpha_{n,t}\neq 0) = \Pi_0 P_{\mathrm{f}_n} + \Pi_1 P_{\mathrm{d}_n}, \quad (3)$$

where the probabilities P_{f_n} and P_{d_n} are

$$P_{\mathrm{f}_n} = \Pr(\xi_n(x_{n,t}) > \theta_n | H_t = 0) = Q\Big(\frac{\theta_n + \mathcal{A}^2/2\sigma_{v_n}^2}{\mathcal{A}/\sigma_{v_n}}\Big),$$
$$P_{\mathrm{d}_n} = \Pr(\xi_n(x_{n,t}) > \theta_n | H_t = 1) = Q\Big(\frac{\theta_n - \mathcal{A}^2/2\sigma_{v_n}^2}{\mathcal{A}/\sigma_{v_n}}\Big). \quad (4)$$

B. Markovian Battery State and Energy Harvesting Models

We assume sensors are equipped with identical batteries of finite size K cells (units), where each cell corresponds to b_u Joules of stored energy. Therefore, each battery is capable of storing at most Kb_u Joules of harvested energy. Let $B_{n,t} \in \{0, 1, ..., K\}$ denote the discrete random process indicating the battery state of sensor n at the beginning slot t. Note that $B_{n,t} = 0$ and $B_{n,t} = K$ represent the empty battery and full battery levels, respectively. Also, $B_{n,t} = k$ implies that the battery is at state k, i.e., k cells of the battery is charged and the amount of stored energy in the battery is kb_u Joules.

Let $\mathcal{E}_{n,t}$ denote the randomly arriving energy units during time slot t at sensor n, where each unit is b_u Joules. We assume $\mathcal{E}_{n,t}$'s are i.i.d. over time slots and across sensors. We model $\mathcal{E}_{n,t}$ as a Poisson random variable with parameter ρ, and probability mass function (pmf) $p_m \triangleq \Pr(\mathcal{E}_{n,t} = m) = e^{-\rho}\rho^m/m!$ for $m = 0, 1, \ldots, \infty$. Note that parameter ρ is the average number of arriving energy units during one time slot at each sensor. Let $\mathcal{S}_{n,t}$ be the number of stored (harvested) energy units in the battery at sensor n during time slot t. Note that the harvested energy $\mathcal{S}_{n,t}$ cannot be used during slot t. Since the battery has a finite capacity of K cells, we have $\mathcal{S}_{n,t} \in \{0, 1, ..., K\}$. Also, $\mathcal{S}_{n,t}$ are i.i.d. over time slots and across sensors. The two random variables $\mathcal{S}_{n,t}$ and $\mathcal{E}_{n,t}$ are related as the following

$$\mathcal{S}_{n,t} = \begin{cases} \mathcal{E}_{n,t}, & \text{if } 0 \le \mathcal{E}_{n,t} \le K-1, \\ K, & \text{if } \mathcal{E}_{n,t} \ge K. \end{cases} \quad (5)$$

Based on (5) we can find the pmf of $\mathcal{S}_{n,t}$ in terms of the pmf of $\mathcal{E}_{n,t}$. Let $q_e \triangleq \Pr(\mathcal{S}_{n,t} = e)$ for $e = 0, 1, \ldots, K$. We have

$$q_e = \begin{cases} p_e, & \text{if } 0 \le e \le K-1, \\ \sum_{m=K}^{\infty} p_m, & \text{if } e = K. \end{cases} \quad (6)$$

The battery state at the beginning of slot $t+1$ depends on the battery state at the beginning of slot t, the harvested energy during slot t, and the transmission symbol $\alpha_{n,t}$, i.e.,

$$B_{n,t+1} = \min\left\{[B_{n,t} + \mathcal{S}_{n,t} - \alpha_{n,t}^2 T_s/b_u]^+, K\right\}, \quad (7)$$

where $[x]^+ = \max\{0, x\}$. Considering the dynamic battery state model in (7) we note that, conditioned on $\mathcal{S}_{n,t}$ and $\alpha_{n,t}$ the value of $B_{n,t+1}$ only depends on the value of $B_{n,t}$. Hence, the process $B_{n,t}$ can be modeled as a Markov chain. Let $\mathbf{\Phi}_{n,t}$ be the probability vector of battery state in slot t

$$\mathbf{\Phi}_{n,t} \triangleq \Big[\Pr(B_{n,t} = 0), \ldots, \Pr(B_{n,t} = K)\Big]^T, \quad (8)$$

where $\Pr(B_{n,t} = k)$ in (8) depends on $B_{n,t-1}$, $\mathcal{S}_{n,t-1}$ and $\alpha_{n,t-1}$. Assuming that the Markov chain is time-homogeneous, we let $\mathbf{\Psi}_n$ be the transition probability matrix of this chain with its (i,j)-th entry $[\mathbf{\Psi}_n]_{i,j} \triangleq \Pr(B_{n,t} = j | B_{n,t-1} = i)$ for $i, j = 0, \ldots, K$. We can express $[\mathbf{\Psi}_n]_{i,j}$ as (10). Since the Markov chain characterized by $\mathbf{\Psi}_n$ is irreducible and aperiodic, there exists a unique steady state distribution, regardless of the initial state [20]. Let $\mathbf{\Phi}_n = [\phi_{n,0}, \phi_{n,1}, ..., \phi_{n,K}]^T$ be the unique steady state probability vector with the entries $\phi_{n,k} = \lim_{t\to\infty} \Pr(B_{n,t} = k)$. This vector satisfies the eigenvalue equation $\mathbf{\Phi}_n = \mathbf{\Phi}_n \mathbf{\Psi}_n$.

C. Markovian Channel Gain Model

During time slot t we assume N sensors send their transmission symbols $\alpha_{n,t}$ simultaneously to the FC, that is equipped with M receive antennas. Let $g_{m,n,t}$ indicate the fading channel gain between sensor n and m-th antenna of the FC during time slot t. The $M \times N$ channel matrix $\boldsymbol{G}_t$ becomes

$$\boldsymbol{G}_t = \begin{bmatrix} g_{1,1,t} & g_{1,2,t} & \cdots & g_{1,N,t} \\ g_{2,1,t} & g_{2,2,t} & \cdots & g_{2,N,t} \\ \vdots & \vdots & \vdots & \vdots \\ g_{M,1,t} & g_{M,2,t} & \cdots & g_{M,N,t} \end{bmatrix}, \quad (9)$$

where $g_{m,n,t}$'s are correlated over time slots, while are independent across sensors and across receive antennas. We define the *channel gain* as the Frobenius norm of $\boldsymbol{G}_t$, i.e., $s_t = ||\boldsymbol{G}_t||_F^2$ [21]. We consider a scalar quantizer at the FC that maps s_t into a point in set $\mathbb{S} = \{\hat{s}_1, \hat{s}_2, ..., \hat{s}_L\}$, which contains L *quantized channel gain* values. The points in set $\mathbb{S}$ can be found such that a certain distortion function is minimized. The FC partitions the positive real line $\mathbb{R}^+$ into L intervals (Voronoi cells of the quantizer) using L quantization thresholds $\{\mu_l\}_{l=1}^L$, where $0 = \mu_1 < \mu_2 < \ldots < \mu_{L-1} < \mu_L = \infty$, and associates interval $\mathcal{I}_l = [\mu_l, \mu_{l+1})$ with point $\hat{s}_l$, i.e., if s_t lies in the interval $\mathcal{I}_l$ then the quantized channel gain $Q(s_t)$ becomes $\hat{s}_l$. We model the time variation of the quantized channel gain using a Markov chain [22]. The Markov chain has L states and the states are the points in set $\mathbb{S}$. To obtain this Markov model, similar to [21], we make the following two assumptions: **(AS1)** The entries of $\boldsymbol{G}_t$ have the Jakes-Clark's correlation function [23], i.e., we have $\mathbb{E}\big[g_{i,j,t}^* \; g_{i,j,t+\tau}\big] = J_0(2\pi f_D \tau), \forall i, j$, where J_0 is Bessel function of zeroth-order and f_D is the maximum Doppler frequency [24]. **(AS2)** Inter-state transitions only occur between adjacent states in the chain. Let $\pi_l = \Pr(Q(s_t) = \hat{s}_l)$ be the steady-state probability of state l of the Markov chain. We have $\pi_l = \int_{\mu_l}^{\mu_{l+1}} f_s(s) ds$, where $f_s(s)$ is the probability density function (pdf) of s_t. Assuming that the elements of $\boldsymbol{G}_t$ are i.i.d and distributed as $\mathcal{CN}(0, 1)$, the channel gain s_t follows a chi-squared distribution with degree of freedom equal to MN. Hence, π_l can be

$$[\mathbf{\Psi}_n]_{i,j} = \widehat{\Pi}_{n,1}\sum_{k=0}^{K}\sum_{l=1}^{L}\pi_l q_k I_{i\to j}(\mathcal{S}_{n,t},\lfloor c_l i\rfloor) + \widehat{\Pi}_{n,0}\sum_{k=0}^{K} q_k I_{i\to j}(\mathcal{S}_{n,t},0),$$

$$\text{where } I_{i\to j}(\mathcal{S}_{n,t},\alpha_{n,t}^2 T_s/b_u) = \begin{cases}1, & \text{if } j = \min\left\{[i+\mathcal{S}_{n,t}-\alpha_{n,t}^2 T_s/b_u]^+, K\right\}.\\ 0, & \text{o.w.}\end{cases} \tag{10}$$

written as

$$\pi_l = \Pr(Q(s_t)=\hat{s}_l) = \sum_{i=0}^{2MN-1}\frac{\exp(-\mu_l)\mu_l^i - \exp(-\mu_{l+1})\mu_{l+1}^i}{i!}.$$

Let Θ be the transition probability matrix of this chain with its (i,j)-th entry $[\Theta]_{i,j} = \Pr(Q(s_t)=\hat{s}_i|Q(s_{t-1})=\hat{s}_j)$. We have

$$[\Theta]_{i,j} = \begin{cases}\frac{\beta(\mu_{l+1}^2)}{\pi_{n,l}}, & i=l+1,\ j=1,...,L-1\\ \frac{\beta(\mu_l^2)}{\pi_{n,l}}, & i=l-1,\ j=2,...,L\\ 1-\frac{\beta(\mu_l^2)}{\pi_{n,l}}-\frac{\beta(\mu_{l+1}^2)}{\pi_{n,l}}, & i=l,\ j=2,...,L-1\\ 1-\frac{\beta(\mu_2^2)}{\pi_{n,1}}, & i=1,\ j=1\\ 1-\frac{\beta(\mu_L^2)}{\pi_{n,L}}, & i=L,\ j=L\\ 0, & \text{O.W.}\end{cases}$$

where β is the level crossing rate of the random process s_t^2 at the level x and is given by [21] $\beta(x) = \frac{\sqrt{2\pi} f_D T_s x^{(Z-1/2)}}{(Z-1)!\exp(x)}$.

D. Transmission Symbol, Received Signals at FC, and Optimal Bayesian Fusion Rule

We consider a simple feedback strategy, in which FC sends the quantized channel gain through a feedback channel to all sensors, every $T>1$ time slots. At time slot t sensor n chooses $\alpha_{n,t}$ according to its *current battery state* k and *the latest available quantized channel gain*, using the following rule

$$\alpha_{n,t}^2 = \begin{cases}0, & \xi_n(x_{n,t})<\theta_n,\\ \lfloor c_1 k\rfloor b_u/T_s, & \xi_n(x_{n,t})\ge\theta_n,\ Q(s_{t'})=\hat{s}_1,\\ \vdots & \qquad\vdots\\ \lfloor c_L k\rfloor b_u/T_s, & \xi_n(x_{n,t})\ge\theta_n,\ Q(s_{t'})=\hat{s}_L,\end{cases} \tag{11}$$

where $\lfloor . \rfloor$ is the floor function, index $t'\in\{t,t-1,...,t-T\}$, and the scale factors $\{c_l\}_{l=1}^L$ are between zero and one and are *our optimization variables*. In each time slot, sensors send their symbols simultaneously to the FC. The received signal at the FC corresponding to time slot t is $\boldsymbol{y}_t = \boldsymbol{G}_t\boldsymbol{\alpha}_t + \boldsymbol{w}_t$, where $\boldsymbol{y}_t = [y_{1,t}, y_{2,t}, ,..., y_{M,t}]^T$, $\boldsymbol{\alpha}_t = [\alpha_{1,t}, \alpha_{2,t}, ,..., \alpha_{N,t}]^T$, $\boldsymbol{w}_t = [w_{1,t}, w_{2,t}, ,..., w_{M,t}]^T$, and $\boldsymbol{w}_t$ is a zero mean complex Gaussian vector with covariance matrix $\boldsymbol{R}$. The FC applies the optimal Bayesian fusion rule $\Gamma_0(.)$ to obtain a global decision $u_{0,t}$ [2]. In particular, we have

$$u_{0,t} = \Gamma_0(\boldsymbol{y}_t) = \begin{cases}1, & \Delta_t>\tau,\\ 0, & \Delta_t<\tau,\end{cases}\quad \Delta_t = \log\left(\frac{f(\boldsymbol{y}_t|H_t=1)}{f(\boldsymbol{y}_t|H_t=0)}\right)$$

where $f(\boldsymbol{y}_t|H_t=h_t)$ is the conditional pdf of $\boldsymbol{y}_t$ and the decision threshold $\tau = \log(\frac{\Pi_0}{\Pi_1})$. From Bayesian perspective, the natural choice to measure the detection performance is the error probability, defined as $P_e = \Pi_0\Pr(\Delta_t>\tau|H_t=0) + \Pi_1\Pr(\Delta_t<\tau|H_t=1)$. However, finding a closed form expression for P_e is mathematically intractable. Instead, we choose the J-divergence between the distributions of the detection statistics at the FC under different hypotheses, as our detection performance metric. This choice allows us to provide a tractable analysis. Given the local thresholds $\{\theta_n\}_{n=1}^N$ in (11) and the channel gain quantizer at the FC, *our problem of optimizing transmit power control strategy reduces to finding the optimal scale factors* $\{c_l\}_{l=1}^L$ *in (11) such that the J-divergence at the FC is maximized, subject to per-sensor average transmit power constraints.*

III. J-Divergence Derivation and Our Constrained Optimization Problem

By definition [5], [7], the J-divergence between two pdfs $\eta_1(x)$ and $\eta_0(x)$, denoted as $J(\eta_1,\eta_0)$, is $J(\eta_1,\eta_0) = D(\eta_1||\eta_0)+D(\eta_0||\eta_1)$, where $D(\eta_i||\eta_j)$ is the non-symmetric Kullback-Leibler (KL) distance between $\eta_i(x)$ and $\eta_j(x)$. The KL distance $D(\eta_i||\eta_j)$ is defined as $D(\eta_i||\eta_j) = \int_{-\infty}^{\infty}\log\left(\frac{\eta_i(x)}{\eta_j(x)}\right)\eta_i(x)dx$. Therefore, we obtain

$$J(\eta_1,\eta_0) = \int_{-\infty}^{\infty}[\eta_1(x)-\eta_0(x)]\log\left(\frac{\eta_1(x)}{\eta_0(x)}\right)dx. \tag{12}$$

In our problem setup, $f(\boldsymbol{y}_t|\boldsymbol{G}_t,H_t=1)$ and $f(\boldsymbol{y}_t|\boldsymbol{G}_t,H_t=0)$ play the role of $\eta_1(x)$ and $\eta_0(x)$, respectively. Given $\boldsymbol{G}_t$ we note that H_t, $\boldsymbol{\alpha_t}$, $\boldsymbol{y}_t$ satisfy the Markov property, i.e., $H_t\to\boldsymbol{\alpha_t}\to\boldsymbol{y}_t$ [5], [7]. This implies that $\boldsymbol{y}_t$ and H_t, given $\boldsymbol{\alpha_t}$, are conditionally independent. Therefore, we can write $f(\boldsymbol{y}_t|\boldsymbol{G}_t,H_t=i) = f(\boldsymbol{y}_t|\boldsymbol{G}_t,\boldsymbol{\alpha_t}=0)\Pr(\boldsymbol{\alpha}_t|H_t=i) + f(\boldsymbol{y}_t|\boldsymbol{G}_t,\boldsymbol{\alpha_t}\neq 0)\Pr(\boldsymbol{\alpha}_t|H_t=i)$ for $i=0,1$. We have

$$f(\boldsymbol{y}_t|\boldsymbol{G}_t,\boldsymbol{\alpha}_t) = \frac{1}{|2\pi\boldsymbol{R}|^{\frac{1}{2}}}\exp[-\frac{1}{2}(\boldsymbol{y}_t-\boldsymbol{G}_t\boldsymbol{\alpha}_t)\boldsymbol{R}^{-1}(\boldsymbol{y}_t-\boldsymbol{G}_t\boldsymbol{\alpha}_t)]$$

Although $f(\boldsymbol{y}_t|\boldsymbol{G}_t,\boldsymbol{\alpha}_t)$ is Gaussian, $f(\boldsymbol{y}_t|\boldsymbol{G}_t,H_t=0), f(\boldsymbol{y}_t|\boldsymbol{G}_t,H_t=1)$ are Gaussian mixtures. Unfortunately, the J-divergence between two Gaussian mixture densities does not have a general closed-form expression. Similar to [5], [7], we approximate the J-divergence between two Gaussian mixture densities by the J-divergence between two Gaussian densities $f^G(\boldsymbol{y}_t|\boldsymbol{G}_t,H_t=i)\sim\mathcal{N}(\boldsymbol{m}_i,\boldsymbol{\Upsilon}_i)$, where the mean and the variance of the approximate distributions are obtained from matching the first and second order moments of the actual and the approximate distributions. For our problem setup, the parameters $\boldsymbol{m}_0,\boldsymbol{m}_1,\boldsymbol{\Upsilon}_0,\boldsymbol{\Upsilon}_1$ become

$$\boldsymbol{m}_0 = \boldsymbol{G}_t\boldsymbol{A}_t\boldsymbol{P}_{\text{f}},\quad \boldsymbol{\Upsilon}_0 = \boldsymbol{R}+\boldsymbol{G}_t\boldsymbol{A}_t\widehat{\boldsymbol{P}}_{\text{f}}\boldsymbol{A}_t^T\boldsymbol{G}_t^T,$$
$$\boldsymbol{m}_1 = \boldsymbol{G}_t\boldsymbol{A}_t\boldsymbol{P}_{\text{d}},\quad \boldsymbol{\Upsilon}_1 = \boldsymbol{R}+\boldsymbol{G}_t\boldsymbol{A}_t\widehat{\boldsymbol{P}}_{\text{d}}\boldsymbol{A}_t^T\boldsymbol{G}_t^T. \tag{13}$$

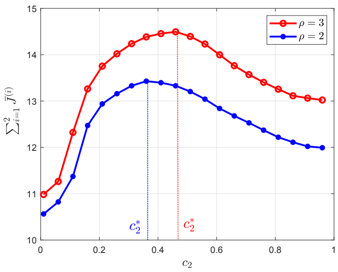

Fig. 2: $M=4$

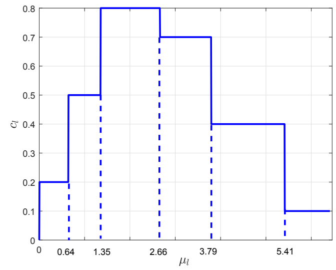

Fig. 3: $M=4, \rho=2$

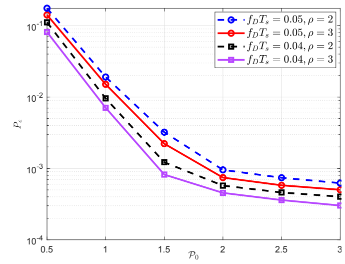

Fig. 4: $M=4$

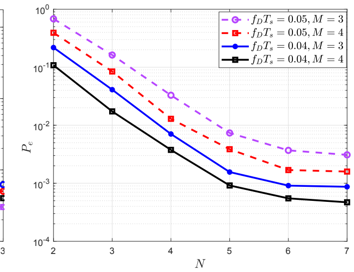

Fig. 5: $\rho=3$

in which $\boldsymbol{A}_t = \text{diag}\{\alpha_{1,t}, ..., \alpha_{N,t}\}$, $\boldsymbol{P}_\text{f} = [P_{\text{f}_1}, ..., P_{\text{f}_N}]^T$, $\boldsymbol{P}_\text{d} = [P_{\text{d}_1}, ..., P_{\text{d}_N}]^T$, $\widehat{\boldsymbol{P}_\text{f}} = \text{diag}\{P_{\text{f}_1}(1-P_{\text{f}_1}), ..., P_{\text{f}_N}(1-P_{\text{f}_N})\}$, and $\widehat{\boldsymbol{P}_\text{d}} = \text{diag}\{P_{\text{d}_1}(1-P_{\text{d}_1}), ..., P_{\text{d}_N}(1-P_{\text{d}_N})\}$. After some algebra, we obtain

$$
\begin{aligned}
&J\big(f^G(\boldsymbol{y}_t|\boldsymbol{G}_t, H_t=1), f^G(\boldsymbol{y}_t|\boldsymbol{G}_t, H_t=0)\big) = \\
&\frac{1}{2}\text{Tr}\big[\boldsymbol{\Upsilon}_0\boldsymbol{\Upsilon}_1^{-1} + \boldsymbol{\Upsilon}_1\boldsymbol{\Upsilon}_0^{-1} \\
&+ (\boldsymbol{\Upsilon}_1^{-1} + \boldsymbol{\Upsilon}_0^{-1})(\boldsymbol{m}_1 - \boldsymbol{m}_0)(\boldsymbol{m}_1 - \boldsymbol{m}_0)^T - \boldsymbol{R}\big]
\end{aligned}
\tag{14}
$$

Note that J in (14) depends on $\boldsymbol{G}_t$, whereas $\alpha_{n,t}^2$ in (11) depends on the quantization interval to which $s_t = ||\boldsymbol{G}_t||_F^2$ belongs. Let $\bar{J}^{(i)} = \mathbb{E}\{J|s_t \in [\mu_i, \mu_{i+1}]\}$ and $\bar{\mathcal{P}}_n^{(i)} = \mathbb{E}\{\alpha_{n,t}^2|s_t \in [\mu_i, \mu_{i+1}]\}$, respectively, denote the expectations of J in (14) and $\alpha_{n,t}^2$ in (11) over s_t, conditioned that $s_t \in [\mu_i, \mu_{i+1}]$. Since $\bar{J}^{(i)}$ does not have a closed-form expression we compute it via Monte Carlo simulation. Using (11) we find $\bar{\mathcal{P}}_n^{(i)} = \widehat{\Pi}_{n,1}\sum_{k=0}^K \phi_{n,k}\pi_i\lfloor c_i k\rfloor$. Our constrained optimization problem of maximizing the J-divergence, subject to per-sensor average transmit power constraints, with respect to the optimization variables $\{c_l\}_{l=1}^L$ in (11) become

$$
\text{(P1)} \qquad \max_{\{c_l\}_{l=1}^L} \sum_{i=1}^L \bar{J}^{(i)}
$$

$$
\text{s.t. } c_l \in [0,1], \forall l, \ \sum_{i=1}^L \bar{\mathcal{P}}_n^{(i)} \leq \mathcal{P}_0, \forall n.
$$

where $\mathcal{P}_0$ is the maximum allowed average transmit power per-sensor. We note that (P1) is *not concave* with respect to the optimization variables. Moreover, the objective function and the constraints in (P1) are *not differentiable* with respect to the optimization variables. Hence, existing gradient-based algorithms for solving non-convex optimization problems cannot be used to solve (P1). We resort to a grid-based search method, which requires L-dimensional search over the search space $[0,1]^L$. Clearly, the accuracy of this solution depends on the resolution of the grid-based search. Suppose the intervals $[0,1]$ is divided into N_c sub-intervals. Therefore, the search space of (P1), denoted as $\mathcal{D}$, consists of $(N_c)^L$ discrete points in the original L-dimensional search space.

Computational complexity of solving (P1): We note that the FC needs to perform two tasks for each point in $\mathcal{D}$: task (i) forming $\boldsymbol{\Psi}_n$ and $\boldsymbol{\Phi}_n$, task (ii) calculating $\bar{J}^{(i)}$ and $\bar{\mathcal{P}}_n^{(i)}$. Our numerical results show that for a fixed $\{c_l\}_{l=1}^L$ the computational complexity of task (i) and task (ii) are $\mathcal{O}(K^{3.2})$ and $\mathcal{O}(M \times N \times K^{2.7})$, respectively. Hence, the computational complexity of solving (P1) is $\mathcal{O}\left((N_c)^L(K^{3.2} + M \times N \times K^{2.7})\right)$. To curb the computational complexity of the grid-based search method, we plan to explore random search algorithms (in which only a randomly chosen subset of the points in $\mathcal{D}$ is searched to find a solution) that have a low-computational complexity and provide a close-to-optimal performance for our future work.

IV. Simulation Results and Concluding Remarks

In our simulations, we let $\boldsymbol{R} = \boldsymbol{I}\sigma_w^2$ and define the SNR corresponding to observation channel as $\text{SNR}_s = 20\log(\mathcal{A}/\sigma_v)$. We let $P_{d_n} = 0.9, \forall n$, $\text{SNR}_s = 3\text{dB}$, $\mathcal{P}_0 = 2\text{mW}$ (except Fig. 4), $N=3$ (except Fig. 5), $\sigma_w^2 = 1, K=5, f_DT_s = 0.05$ (except Fig. 4 and 5), and $T=10$ (feedback is sent every 10 slots). For $L=2$ the optimization variables are $\{c_1, c_2\}$. Fig. 2 illustrates the objective function $\bar{J}^{(1)} + \bar{J}^{(2)}$ versus the scale factor c_2 given $c_1 = 0.5$. We observe that the objective function is not a concave function of c_2. Still there exists a point, denoted as c_2^*, at which the function attains its maximum. Starting from small values of c_2, as c_2 increases (until it reaches c_2^*), the function value increases, because the harvested energy can recharge the battery and can yield more power for data transmission. However, when c_2 exceeds c_2^*, the harvested and stored energy cannot support the data transmission and the function value decreases. Fig. 3 depicts the optimized $\{c_l\}$'s versus the quantization thresholds $\{\mu_l\}$'s for $L=6$. We note that, as l increases (i.e., channel gain s_t increases), c_l first increases and then decreases. Considering (11) this implies that, given the battery state k, as s_t increases $\alpha_{n,t}^2$ first increases and then decreases. Fig. 4 shows the error probability P_e versus $\mathcal{P}_0$, as f_DT_s and ρ vary for $L=4$. Given the pair (f_DT_s, ρ), as $\mathcal{P}_0$ increases P_e decreases. Also, P_e decreases when (i) given the pair ($\mathcal{P}_0$, ρ), f_DT_s decreases; (ii) given the pair ($\mathcal{P}_0$, f_DT_s), ρ increases. Fig. 5 depicts P_e versus N as f_DT_s and M vary for $L=4$. Given the pair (f_DT_s, M), as N increases P_e reduces, until it reaches an error floor. This is because for larger N values, P_e becomes limited by the communication channel noise σ_w^2. Furthermore, we notice that P_e decreases when (i) given the pair (N, M), f_DT_s decreases; (ii) given the pair (N, f_DT_s), M increases.

V. Acknowledgement

The work of Azadeh Vosoughi is supported by U.S. National Science Foundation award 16226A19.

References

[1] D. Ciuonzo, P. S. Rossi, and P. K. Varshney, "Distributed detection in wireless sensor networks under multiplicative fading via generalized score tests," *IEEE Internet of Things Journal*, vol. 8, no. 11, pp. 9059–9071, 2021.

[2] H. R. Ahmadi and A. Vosoughi, "Distributed detection with adaptive topology and nonideal communication channels," *IEEE transactions on signal processing*, vol. 59, no. 6, pp. 2857–2874, 2011.

[3] B. Chen, L. Tong, and P. K. Varshney, "Channel-aware distributed detection in wireless sensor networks," *IEEE Signal Processing Magazine*, vol. 23, no. 4, pp. 16–26, 2006.

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

[5] Z. Hajibabaei, A. Vosoughi, and N. Mastronarde, "Optimal power allocation for M-ary distributed detection in the presence of channel uncertainty," *Signal Processing*, vol. 169, p. 107400, 2020.

[6] H. R. Ahmadi and A. Vosoughi, "Impact of channel estimation error on decentralized detection in bandwidth constrained wireless sensor networks," in *IEEE Military Communications Conference (MILCOM)*. IEEE, 2008, pp. 1–7.

[7] X. Zhang, H. V. Poor, and M. Chiang, "Optimal power allocation for distributed detection over MIMO channels in wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 56, no. 9, pp. 4124–4140, Sep. 2008.

[8] N. Maleki and A. Vosoughi, "On bandwidth constrained distributed detection of a known signal in correlated gaussian noise," *IEEE Transactions on Vehicular Technology*, 2020.

[9] N. Maleki, A. Vosoughi, and N. Rahnavard, "Distributed binary detection over fading channels: Cooperative and parallel architectures," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 9, pp. 7090–7109, 2015.

[10] A. Sani and A. Vosoughi, "Bandwidth and power constrained distributed vector estimation in wireless sensor networks," in *IEEE Military Communications Conference (MILCOM)*, 2015.

[11] ——, "On distributed linear estimation with observation model uncertainties," *IEEE Transactions on signal processing*, vol. 66, no. 12, pp. 3212–3227, 2018.

[12] M. Shirazi and A. Vosoughi, "On distributed estimation in hierarchical power constrained wireless sensor networks," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 6, pp. 442–459, 2020.

[13] N. Maleki and A. Vosoughi, "Channel-aware m-ary distributed detection: Optimal and suboptimal fusion rules," in *IEEE Statistical Signal Processing Workshop (SSP)*, 2012.

[14] Z. Esmaeilbeig, K. V. Mishra, and M. Soltanalian, "IRS-aided radar: Enhanced target parameter estimation via intelligent reflecting surfaces," in *IEEE 12th Sensor Array and Multichannel Signal Processing Workshop (SAM)*, 2022, pp. 286–290.

[15] S. Sudevalayam and P. Kulkarni, "Energy harvesting sensor nodes: Survey and implications," *IEEE Communications Surveys Tutorials*, vol. 13, no. 3, pp. 443–461, Third 2011.

[16] G. Ardeshiri and A. Vosoughi, "On adaptive transmission for distributed detection in energy harvesting wireless sensor networks with limited fusion center feedback," *IEEE Transactions on Green Communications and Networking*, vol. 6, no. 3, pp. 1764–1779, 2022.

[17] G. Ardeshiri, H. Yazdani, and A. Vosoughi, "Power adaptation for distributed detection in energy harvesting wsns with finite-capacity battery," in *IEEE Global Communications Conference (GLOBECOM)*. IEEE, 2019, pp. 1–6.

[18] G. Ardeshiri, H. Yazdani, and A. Vosoughi, "Optimal local thresholds for distributed detection in energy harvesting wireless sensor networks," in *2018 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, Nov 2018, pp. 813–817.

[19] G. Ardeshiri and A. Vosoughi, "Learning-based distributed detection with energy harvesting," in *55th Asilomar Conference on Signals, Systems, and Computers*, 2021, pp. 747–751.

[20] J. F. Shortle, J. M. Thompson, D. Gross, and C. M. Harris, *Fundamentals of queueing theory*. John Wiley & Sons, 2018, vol. 399.

[21] K. Huang, B. Mondal, W. Heath, and J. Andrews, "Markov models for limited feedback mimo systems," in *2006 IEEE International Conference on Acoustics Speech and Signal Processing Proceedings*, vol. 4, 2006, pp. IV–IV.

[22] S. Akoum and R. W. Heath, "Limited feedback beamforming for temporally correlated mimo channels with other cell interference," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2010, pp. 3054–3057.

[23] W. C. Jakes and D. C. Cox, *Microwave mobile communications*. Wiley-IEEE press, 1994.

[24] Hong Shen Wang and N. Moayeri, "Finite-state markov channel-a useful model for radio communication channels," *IEEE Transactions on Vehicular Technology*, vol. 44, no. 1, pp. 163–171, Feb 1995.