INTERNATIONAL JOURNAL OF GENERAL SYSTEMS, 2018
VOL. 47, NO. 5, 529–542
https://doi.org/10.1080/03081079.2018.1455192



# DECHADE: DEtecting slight Changes with HArd DEcisions in Wireless Sensor Networks

D. Ciuonzo[a] and P. Salvo Rossi[b]

[a]Network Measurement and Monitoring (NM2) s.r.l., Naples, Italy; [b]Kongsberg Digital, Trondheim, Norway

**ABSTRACT**

This paper focuses on the problem of change detection through a Wireless Sensor Network (WSN) whose nodes report only binary decisions (on the presence/absence of a certain event to be monitored), due to bandwidth/energy constraints. The resulting problem can be modelled as testing the equality of samples drawn from independent Bernoulli probability mass functions, when the bit probabilities under both hypotheses are not known. Both One-Sided (OS) and Two-Sided (TS) tests are considered, with reference to: (i) identical bit probability (a homogeneous scenario), (ii) different per-sensor bit probabilities (a non-homogeneous scenario) and (iii) regions with identical bit probability (a block-homogeneous scenario) for the observed samples. The goal is to provide a systematic framework collecting a plethora of viable detectors (designed via theoretically founded criteria) which can be used for each instance of the problem. Finally, verification of the derived detectors in two relevant WSN-related problems is provided to show the appeal of the proposed framework.

**ARTICLE HISTORY**
Received 29 October 2017
Accepted 15 February 2018

**KEYWORDS**
Change detection; Data Fusion; Internet of Things; Wireless Sensor Networks

## 1. Introduction

Hypothesis testing from Bernoulli samples[1] is of paramount importance in several Wireless Sensor Networks (WSNs) applications (Niu and Varshney 2008; Ciuonzo et al. 2013; Ciuonzo and Salvo Rossi 2014; Ciuonzo, De Maio, and Salvo Rossi 2015). One relevant example is the *Decentralized Change Detection* (DCD) problem (He, Ben-David, and Tong 2006), i.e. a binary hypothesis test where a Fusion Centre (FC) is employed to detect a *triggering event*, namely a change in a Phenomenon Of Interest (POI, e.g. a target, a source) between consecutive discrete-time instants, based on binary decisions from a WSN (e.g. presence/absence of the POI being monitored (Varshney 1996)).

The design of practical fusion rules for this problem has one of its main drivers in resource-limited (bandwidth, energy, computation) WSNs, i.e. a major player in the emerging Internet of Things (IoT) paradigm. Relevant applications include anomaly detection, environmental change monitoring and uncooperative target intent inference. The use of WSNs is mainly motivated by the (spatial) diversity they can offer, whose effective data processing may translate into increased robustness of the system. To the best of our knowledge, although many works have dealt with DCD (Patwari, Hero, and Sadler

**CONTACT** D. Ciuonzo domenico.ciuonzo@ieee.org

2003; He, Ben-David, and Tong 2006; Banerjee and Veeravalli 2015; Faivishevsky 2016), none of them has provided a comprehensive study of fusion rules design, capitalizing on severely constrained WSNs and based on composite hypothesis testing tools, so as to be *non-parametric but theoretically grounded*.

In Patwari, Hero, and Sadler (2003), a CUSUM test for hierarchical distributed detection based on censoring sensors is proposed. Unfortunately, *full measurement reporting is assumed* and, equally important, the PDF of the *pre/post-change* is assumed to be *known*. In Banerjee and Veeravalli (2015), DCD is considered and minimax solutions are proposed therein. The approach both allows for controlling the cost of sensors' observations taken and the communication cost between the sensors and the FC. Unfortunately, again the PDF of the *pre/post-change* is assumed to be *known*. Finally, in Faivishevsky (2016), DCD is (partially) explored with application to IoT, based on a per-sensor information-theoretic multivariate CD method originating from $K$-nearest neighbour estimation. The proposed method is favourably tested on both simulated an real data; however, the proposed fusion technique requires availability of full-precision statistics at the FC.

The closest work to ours is represented by He, Ben-David, and Tong (2006), where DCD via one-bit decisions is considered *only* under Two-Sided (TS) testing. A non-parametric DCD framework is there developed, with many detection/estimation algorithms developed based on the Vapnik–Chervonenkis theory. Unfortunately, though intuitive, the design of the considered rules has no theoretical motivation.

The aim of this work is threefold:

- providing a detailed overview of the possible alternatives which can be tackled for the three composite hypothesis tests later specified in Equations (1), (2) and (3), since the optimum Log-Likelihood Ratio (LLR) *cannot be implemented*, due to presence of unknown parameters (Kay 1998);
- investigating the existence of the Uniformly Most Powerful (UMP) test and pointing out that *it does not exist* in all the cases considered;
- deriving the Generalized Likelihood Ratio Test (GLRT), the Rao Test, the Wald Test and the Locally Most (Mean) Powerful Test (LM(M)PT) as viable decision strategies (investigating possible coincidence and/or statistical equivalence).

In many relevant cases, the sensor bit probability is linked to the *detection rate* of a POI to be revealed/monitored, and a change can be modelled as a (positive/negative) shift of the bit probability (Ciuonzo, Salvo Rossi, and Willett 2017). In other cases, the change can be related to an abrupt increase/decrease of the concentration of a certain environmental property (e.g. temperature, received power, (Ciuonzo and Salvo Rossi 2017)). Hence, for sake of completeness, we will focus on both One-Sided (OS) and Two-Sided (TS) testing (Kay 1998). Additionally, we will analyse three different scenarios, shown in Figure 1, as for the change relationship with the spatial dimension: (i) Homogeneous (H), (ii) Non-Homogeneous (NH) and (iii) Block-Homogeneous (BH).

It is worth mentioning that other practical impairments affecting WSNs, such as delayed and/or missing measurements, are not considered in this work. Possible frameworks for including such phenomena into WSN design are found in Basin, Shi, and Calderon-Alvarez (2010) and Caballero-Águila, Hermoso-Carazo, and Linares-Pérez (2015).

The rest of the manuscript is organized as follows: the system model and corresponding problem formulation are presented in Section 2; in Section 3, we derive and discuss the

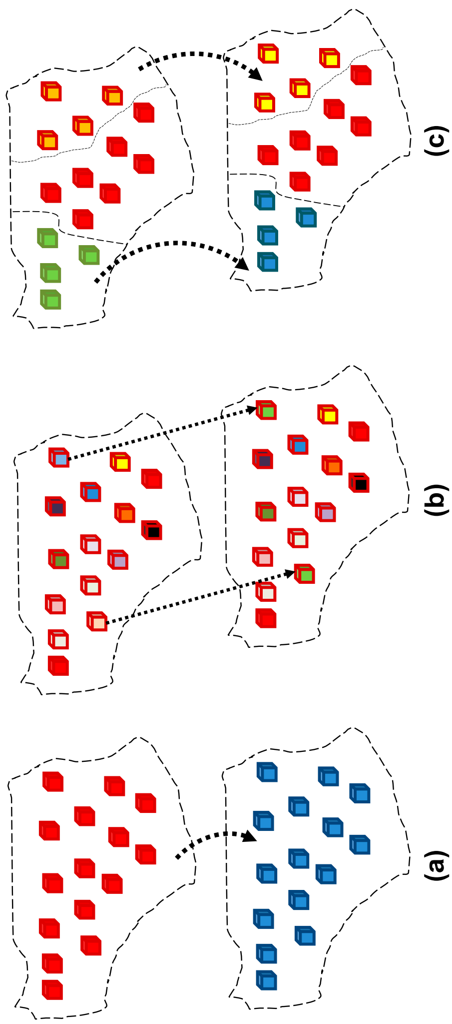



**Figure 1.** DCD in WSNs with different *spatial* assumptions: (a) H, (b) NH and (c) BH.

main results, while numerical results to assess the DECHADE framework are provided in Section 4. Finally, concluding remarks and sparks for future research are given in Section 5.

## 2. System model

We consider a WSN with $K$ sensors monitoring a POI and reporting decisions to a FC for DCD and arranged in a star topology.[2] The null (resp. alternative) hypothesis $\mathcal{H}_0$ (resp. $\mathcal{H}_1$) represents the *absence of change* (resp. an *occurred change*) of the POI. We denote $y_k[i]$ the local decision by the $k$th sensor at the $i$th discrete time, which (under $\mathcal{H}_j$) is characterized by the bit probability $P(y_k[i]=1|\mathcal{H}_j)$.

The WSN model is entirely specified via the joint multivariate PMF under $\mathcal{H}_j$, denoted $P(\boldsymbol{y}[i],\, \boldsymbol{y}[i+1]\,|\,\mathcal{H}_j)$, where $\boldsymbol{y}[\ell] \triangleq \left[y_1[\ell] \cdots y_K[\ell]\right]^T$ collects all data at the $\ell$th discrete time. For simplicity, we assume conditionally independent observations in *both* space and time, i.e. $P(\boldsymbol{y}[i],\, \boldsymbol{y}[i+1]|\mathcal{H}_j) = \prod_{k=1}^{K}\left\{P(y_k[i]|\mathcal{H}_j)\, P(y_k[i+1]|\mathcal{H}_j)\right\}$. More specifically, when $\mathcal{H}_0$ holds (i.e. no change), the $k$th sensor is characterized by the *same* bit probability at both $i$th and $(i+1)$th discrete times, which is denoted $\pi_k \triangleq P\left(y_k[i]=1|\mathcal{H}_0\right) = P(y_k[i+1]=1|\mathcal{H}_0)$. On the other hand, when $\mathcal{H}_1$ holds (i.e. a change occurs), the $k$th sensor is characterized by two *different* bit probabilities denoted $\pi_k \triangleq P\left(y_k[i]=1|\mathcal{H}_1\right)$ and $\varphi_k \triangleq P\left(y_k[i+1]=1|\mathcal{H}_1\right)$, *before* and *after* the change, respectively. In this work, aiming at designing *non-parametric* fusion rules, we assume that both $\pi_k$ and $\varphi_k$ are *unknown* and *deterministic* parameters (Kay 1998).

The three problems corresponding to the scenarios introduced in Section 1 are

$$\begin{cases} \mathcal{H}_0 \;:\; y_k[i] \sim \mathcal{B}(\pi), & y_k[i+1] \sim \mathcal{B}(\pi), \\ \mathcal{H}_1 \;:\; y_k[i] \sim \mathcal{B}(\pi), & y_k[i+1] \sim \mathcal{B}(\varphi), \end{cases} \quad k \in \{1,\ldots K\} \tag{1}$$

in the case of a *H scenario*, and

$$\begin{cases} \mathcal{H}_0 \;:\; y_k[i] \sim \mathcal{B}(\pi_k), & y_k[i+1] \sim \mathcal{B}(\pi_k), \\ \mathcal{H}_1 \;:\; y_k[i] \sim \mathcal{B}(\pi_k), & y_k[i+1] \sim \mathcal{B}(\varphi_k), \end{cases} \quad k \in \{1,\ldots K\} \tag{2}$$

in the case of a *NH scenario*. Finally, in the case of a *BH* scenario, the WSN is subdivided in $M$ spatially homogeneous regions $\mathcal{R}_m$ ($m=1,\ldots M$) and the $K_m$ sensors within each region $\mathcal{R}_m$ are associated to the pair $(\pi_m, \varphi_m)$. In other terms, sensors within $\mathcal{R}_m$ are solely characterized by $\pi_m$ when $\mathcal{H}_0$ holds, whereas when $\mathcal{H}_1$ is in force are characterized by $\pi_m$ and $\varphi_m$ *before* and *after* the change, respectively. The hypothesis test for a BH scenario is then:

$$\begin{cases} \mathcal{H}_0 \;:\; y_n[i] \sim \mathcal{B}(\pi_m), & y_n[i+1] \sim \mathcal{B}(\pi_m), \\ \mathcal{H}_1 \;:\; y_n[i] \sim \mathcal{B}(\pi_m), & y_n[i+1] \sim \mathcal{B}(\varphi_m), \end{cases} \quad n \in \mathcal{R}_m,\; m \in \{1,\ldots M\} \tag{3}$$

Henceforth, the PMF of the decisions when $\mathcal{H}_0$ (resp. $\mathcal{H}_1$) holds will be also denoted $P_0(\boldsymbol{y}[i],\, \boldsymbol{y}[i+1]\,;\,\cdot)$ (resp. $P_1(\boldsymbol{y}[i],\, \boldsymbol{y}[i+1]\,;\,\cdot)$) so as to *underline* the unknown parameters. Furthermore, we remark that the hypothesis test in a H scenario can be restated in terms of *relevant signal parameter* $\Delta \triangleq (\varphi - \pi)$, as follows:

$$\text{(TS)} \quad \begin{cases} \mathcal{H}_0: & \pi,\ \Delta = \Delta_0 \\ \mathcal{H}_1: & \pi,\ \Delta \neq \Delta_0 \end{cases}, \qquad \text{(OS)} \quad \begin{cases} \mathcal{H}_0: & \pi,\ \Delta = \Delta_0 \\ \mathcal{H}_1: & \pi,\ \Delta > \Delta_0 \end{cases}, \tag{4}$$

where $\Delta_0 \triangleq 0$. A similar formulation holds for a NH (resp. BH) scenario when considering $\boldsymbol{\pi} \triangleq \left[\pi_1 \cdots \pi_K\right]^T$ (resp. $\boldsymbol{\pi} \triangleq \left[\pi_1 \cdots \pi_M\right]^T$) as the nuisance parameter vector and $\boldsymbol{\Delta} \triangleq \left[(\varphi_1 - \pi_1) \cdots (\varphi_K - \pi_K)\right]^T$ (resp. $\boldsymbol{\Delta} \triangleq \left[(\varphi_1 - \pi_1) \cdots (\varphi_M - \pi_M)\right]^T$) as the *relevant signal parameter vector*. Accordingly, for TS testing, $\mathcal{H}_1$ corresponds to $\boldsymbol{\Delta} \neq \mathbf{0}_K$ (resp. $\boldsymbol{\Delta} \neq \mathbf{0}_M$), whereas $\mathcal{H}_0$ implies $\boldsymbol{\Delta} = \mathbf{0}_K$ (resp. $\boldsymbol{\Delta} = \mathbf{0}_M$); while, for OS testing, $\mathcal{H}_1$ corresponds to the (generalized) inequality $\boldsymbol{\Delta} \geq .\mathbf{0}_K$ (resp. $\boldsymbol{\Delta} \geq .\mathbf{0}_M$), whereas $\mathcal{H}_0$ implies $\boldsymbol{\Delta} = \mathbf{0}_K$ (resp. $\boldsymbol{\Delta} = \mathbf{0}_M$). Such (equivalent) formulation allows to separate naturally the relevant signal parameters (viz. $\Delta$ and $\boldsymbol{\Delta}$ for H and NH/BH scenarios, respectively) from the nuisance one (viz. $\pi$ and $\boldsymbol{\pi}$ for H and NH/BH scenarios, respectively) and will be frequently exploited in the following.

The goal of a DCD system design is a decision statistics in the form $\Lambda(\boldsymbol{y}[i],\ \boldsymbol{y}[i+1])$ to be compared with a threshold $\gamma_{\text{fc}}$ for declaring $\mathcal{H}_1$ (resp. $\mathcal{H}_0$) if the statistics is larger (resp. smaller). System performance is evaluated in terms of false-alarm probability ($P_{\text{F,fc}} \triangleq \Pr\{\Lambda > \gamma_{\text{fc}} | \mathcal{H}_0\}$) and detection probability ($P_{\text{D,fc}} \triangleq \Pr\{\Lambda > \gamma_{\text{fc}} | \mathcal{H}_1\}$).

## 3. Taxonomy of decision rules

In this section, we develop all the explicit forms of the considered rules for the DCD under investigation. A complete summary of the decision statistics within DECHADE is reported in Table 1 for readers' convenience.[3] Each one of them represents a suitable alternative (to the unfeasible LLR) for implementing a practical test aiming at DCD.

### 3.1. LLR and UMP test

We start from investigating the LLR (and UMP test existence) in the simple H scenario. Indeed, according to Neyman–Pearson criterion, the latter is derived as (Kay 1998):

$$\begin{aligned} \Lambda_{\text{LLR}} &\triangleq \ln \left[ \frac{P_1(\boldsymbol{y}[i],\ \boldsymbol{y}[i+1];\ \pi, \varphi)}{P_0(\boldsymbol{y}[i],\ \boldsymbol{y}[i+1];\ \pi)} \right] \\ &= \sum_{k=1}^{K} \left\{ \ln \left[ \frac{P(y_k[i];\pi)}{P(y_k[i];\pi)} \right] + \ln \left[ \frac{P_1(y_k[i+1];\ \varphi)}{P_0(y_k[i+1];\pi)} \right] \right\} \\ &= c[i+1]\ \ln\left[\varphi/\pi\right] + (K - c[i+1])\ \ln\left[(1-\varphi)/(1-\pi)\right] \end{aligned} \tag{5}$$

where $c[\ell] \triangleq \sum_{k=1}^{K} y_k[\ell]$ (i.e. the counting sum at $\ell$th time instant) and can be shown to be statistically equivalent to $\Lambda_{\text{LLR}} \propto \ln\left[\frac{\varphi\,(1-\pi)}{\pi\,(1-\varphi)}\right]\ c[i+1]$. The last expression apparently depends on the *unknown pair* $(\pi, \varphi)$. We now analyse TS and OS testing, separately. First, it is easily understood that the *UMP test does not exist* when testing a TS alternative, based on the log term (since it can assume either positive or negative values depending on $\varphi$ and $\pi$). Secondly, in OS testing, the condition $\varphi > \pi$ automatically implies positivity of log term, that is $\Lambda_{\mathrm{LLR}} \propto c[i+1]$. Nonetheless, the *counting rule* at $(i+1)$th time instant is not UMP in the considered DCD setup (as opposed to the simpler decision fusion problem, see e.g. Varshney (1996)), as any error-vanishing threshold $\gamma_{\mathrm{fc}}$ for $c[i+1]$ would (even in asymptotic sense) depend on the *unknown* $\pi$.

Similarly, in the NH case, the LLR has the following closed-form:

$$\Lambda_{\mathrm{LLR}} \triangleq \ln\left[P_1(\boldsymbol{y}[i],\ \boldsymbol{y}[i+1]\,;\ \boldsymbol{\pi},\boldsymbol{\varphi})\,/\,P_0(\boldsymbol{y}[i],\ \boldsymbol{y}[i+1]\,;\ \boldsymbol{\pi})\right] \tag{6}$$

$$= \sum_{k=1}^{K}\left\{y_k[i+1]\ln\left[\varphi_k/\pi_k\right] + (1-y_k[i+1])\ln\left[(1-\varphi_k)/(1-\pi_k)\right]\right\}. \tag{7}$$

Clearly, the LLR cannot be evaluated since the pairs $(\pi_k,\varphi_k)$ are *unknown* and statistically equivalent tests cannot be drawn. Then, it is apparent that for both TS and OS testing the *UMP test does not exist*. For such a reason, *evaluation and comparison of viable decision rules is relevant and will be the object of the remainder of this manuscript*. Indeed, analogous considerations as the NH scenario hold for the BH case, whose LLR is similarly expressed as

$$\Lambda_{\mathrm{LLR}} \triangleq \ln\left[P_1(\boldsymbol{y}[i],\ \boldsymbol{y}[i+1]\,;\ \boldsymbol{\pi},\boldsymbol{\varphi})\,/\,P_0(\boldsymbol{y}[i],\ \boldsymbol{y}[i+1]\,;\ \boldsymbol{\pi})\right] \tag{8}$$

$$= \sum_{m=1}^{M}\left\{\bar{c}_m[i+1]\ln\left[\varphi_m/\pi_m\right] + (K_m-\bar{c}_m[i+1])\ln\left[(1-\varphi_m)/(1-\pi_m)\right]\right\}, \tag{9}$$

where $\bar{c}_m[\ell] \triangleq \sum_{k\in\mathcal{R}_m} y_k[\ell]$, $\ell\in\{i,i+1\}$, i.e. the *counting sum* of all the sensors belonging to $\mathcal{R}_m$, depending on the region-specific pair $(\pi_m,\varphi_m)$.

### *3.2. GLRT*

The GLRT is widely used to devise decision rules in composite hypothesis testing (Kay 1998). In the present DCD problem, the GLR for the H *scenario* is given by

$$\Lambda_{\mathrm{G}} \triangleq \ln\left[\frac{\max_{(\pi,\varphi)}P_1(\boldsymbol{y}[i],\boldsymbol{y}[i+1];\pi,\varphi)}{\max_{(\pi)}P_0(\boldsymbol{y}[i],\boldsymbol{y}[i+1];\pi)}\right], \tag{10}$$

and assumes the explicit expression (after simple manipulations):

$$\Lambda_{\mathrm{G}} = \begin{cases} K\,\left\{\mathcal{D}_{\mathrm{KL}}(\mathcal{B}(\hat{\pi}_1)||\mathcal{B}(\hat{\pi}_0)) + \mathcal{D}_{\mathrm{KL}}(\mathcal{B}(\widehat{\varphi}^{+})||\mathcal{B}(\hat{\pi}_0))\right\} & \text{(OS)} \\ K\,\left\{\mathcal{D}_{\mathrm{KL}}(\mathcal{B}(\hat{\pi}_1)||\mathcal{B}(\hat{\pi}_0)) + \mathcal{D}_{\mathrm{KL}}(\mathcal{B}(\widehat{\varphi})||\mathcal{B}(\hat{\pi}_0))\right\} & \text{(TS)} \end{cases}; \tag{11}$$

where $\hat{\pi}_0 \triangleq (c[i]+c[i+1])\,/\,2K$ (i.e. the ML estimate of $\pi$ under $\mathcal{H}_0$), $\hat{\pi}_1 \triangleq c[i]/K$ (i.e. the ML estimate of $\pi$ under $\mathcal{H}_1$), $\hat{\varphi} \triangleq c[i+1]/K$ (i.e. the ML estimate of $\varphi$ in TS testing) and $\hat{\varphi}^{+} \triangleq \max\{\hat{\pi}_1,\hat{\varphi}\}$ (the ML estimate of $\varphi$ under OS constraint).

It is worth mentioning that an intuitive generalization of the well-known test proposed in Hoeffding (1965) arises in TS testing, i.e. a threshold test based on the sample KL divergences between the (estimated) Bernoulli PMFs under the change hypothesis (i.e. $\mathcal{B}(\hat{\pi}_1)$ and $\mathcal{B}(\hat{\varphi})$) and the (estimated) Bernoulli PMF under no change assumption ($\mathcal{B}(\hat{\pi}_0)$). Similarly, in OS testing, a *restricted version* for $\hat{\varphi} > \hat{\pi}_1$ is obtained.

Secondly, in the NH scenario, the GLR requires the following multi-parameter maximization:

$$\Lambda_{\mathrm{G}} \triangleq \sum_{k=1}^{K} \ln \left[ \frac{\max_{(\pi_k,\varphi_k)} P_1(y_k[i], y_k[i+1];\ \pi_k, \varphi_k)}{\max_{(\pi_k)} P_0(y_k[i], y_k[i+1];\ \pi_k)} \right]. \tag{12}$$

It is not difficult to show that the ML estimate of $\pi_k$ under $\mathcal{H}_0$ equals $\hat{\pi}_{k,0} \triangleq (y_k[i] + y_k[i+1])/2$, whereas the ML estimate of the pair $(\pi_k, \varphi_k)$ under $\mathcal{H}_1$ is $\left(\hat{\pi}_{k,1}, \hat{\varphi}_k\right) \triangleq \left(y_k[i],\ y_k[i] \vee y_k[i+1]\right)$ (resp. $\left(\hat{\pi}_{k,1}, \hat{\varphi}_k\right) \triangleq (y_k[i],\ y_k[i+1])$ for OS (resp. TS) testing. Substitution into (12) leads to

$$\Lambda_{\mathrm{G}} = 2 \cdot \ln 2 \cdot \sum_{k=1}^{K} \left\{ y_k[i] \oplus y_k[i+1] \right\}, \tag{13}$$

in *both testing cases*.

Finally, in the BH case, a similar multi-parameter optimization is required as the H scenario, exploiting the block assumptions of this scenario, that is:

$$\Lambda_{\mathrm{G}} \triangleq \sum_{m=1}^{M} \ln \left[ \frac{\max_{(\pi_m,\varphi_m)} P_1(\bar{c}_m[i],\ \bar{c}_m[i+1];\ \pi_m, \varphi_m)}{\max_{(\pi_m)} P_0(\bar{c}_m[i],\ \bar{c}_m[i+1];\ \pi_m)} \right]. \tag{14}$$

In this case, the GLR assumes the explicit expression:

$$\Lambda_{\mathrm{G}} = \begin{cases} K \sum_{m=1}^{M} \rho_m \left\{ \mathcal{D}_{\mathrm{KL}}(\mathcal{B}(\hat{\pi}_{m,1}) || \mathcal{B}(\hat{\pi}_{m,0})) + \mathcal{D}_{\mathrm{KL}}(\mathcal{B}\left(\hat{\varphi}_m^{+}\right) || \mathcal{B}(\hat{\pi}_{m,0})) \right\} & \text{(OS)} \\ K \sum_{m=1}^{M} \rho_m \left\{ \mathcal{D}_{\mathrm{KL}}(\mathcal{B}(\hat{\pi}_{m,1}) || \mathcal{B}(\hat{\pi}_{m,0})) + \mathcal{D}_{\mathrm{KL}}(\mathcal{B}(\hat{\varphi}_m) || \mathcal{B}(\hat{\pi}_{m,1})) \right\} & \text{(TS)} \end{cases}; \tag{15}$$

where we have employed the definitions $\rho_m \triangleq (K_m/K)$ (the fraction of sensors within $\mathcal{R}_m$), $\hat{\pi}_{m,0} \triangleq (\bar{c}_m[i] + \bar{c}_m[i+1]) \,/\, 2K_m$ (the ML estimate of $\pi_m$ under $\mathcal{H}_0$), $\hat{\pi}_{m.1} \triangleq \bar{c}_m[i]/K_m$ (the ML estimate of $\pi_m$ under $\mathcal{H}_1$), $\hat{\varphi}_m \triangleq \bar{c}_m[i+1]/K_m$ (the ML estimate of $\varphi_m$ in TS testing) and $\hat{\varphi}_m^{+} \triangleq \max\{\hat{\pi}_{m,1}, \hat{\varphi}_m\}$ (the ML estimate of $\varphi_m$ under OS constraint). Interestingly in BH case, the result in Equation (15) merely corresponds to the sum of GLR statistics pertaining to all the regions $\mathcal{R}_m$, $m = 1, \ldots, M$, modelling the monitored area.

### 3.3. Rao test

The Rao test is another well-known test widely applied in TS testing problems, having the same asymptotic performance as the GLRT (Kay 1998). In the *H* scenario, the Rao (score) test is evaluated as (Cressie 1978)

$$\Lambda_{\mathrm{R}} \triangleq \left. \left( \frac{\partial \ln \left[ P_1(\boldsymbol{y}[i],\ \boldsymbol{y}[i+1];\ \Delta, \pi) \right]}{\partial \Delta} \right)^2 \right|_{(\Delta_0, \hat{\pi}_0)} \left[ \mathcal{I}(\Delta_0, \hat{\pi}_0)^{-1} \right]_{\Delta,\Delta} = \frac{2K \left( \hat{\varphi} - \hat{\pi}_0 \right)^2}{\hat{\pi}_0 (1 - \hat{\pi}_0)} \tag{16}$$

where $\boldsymbol{\theta} \triangleq \left[ \Delta\ \pi \right]^T$, $\mathcal{I}(\Delta, \pi) \triangleq \mathbb{E} \left\{ \frac{\partial P_1(\boldsymbol{y}[i], \boldsymbol{y}[i+1];\, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{\partial P_1(\boldsymbol{y}[i], \boldsymbol{y}[i+1];\, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}^T} \right\}$ denotes the Fisher Information Matrix (FIM) and $\left[ \mathcal{I}(\Delta, \pi)^{-1} \right]_{\Delta,\Delta}$ indicates the *scalar* obtained by selecting from the FIM inverse only the element corresponding to the parameter $\Delta$. Similarly, the Rao test in the *NH* scenario is evaluated as

$$\Lambda_{\mathrm{R}} \triangleq \sum_{k=1}^{K} \left( \frac{\partial \ln \left[ P_1(y_k[i], y_k[i+1]; \Delta_k, \pi_k) \right]}{\partial \Delta_k} \right)^2 \Bigg|_{(\Delta_0, \hat{\pi}_{k,0})} \left[ \mathcal{I}_k(\Delta_0, \hat{\pi}_{k,0})^{-1} \right]_{\Delta_k, \Delta_k}$$
$$= \sum_{k=1}^{K} 2 \, \frac{\left[ y_k[i+1] - (y_k[i+1] + y_k[i])/2 \right]^2}{(y_k[i+1] + y_k[i])/2 \left[ 1 - (y_k[i+1] + y_k[i])/2 \right]} = 2 \sum_{k=1}^{K} \left\{ y_k[i] \oplus y_k[i+1] \right\} \tag{17}$$

where $\boldsymbol{\theta}_k \triangleq \left[ \Delta_k \; \pi_k \right]^T$, $\mathcal{I}_k(\Delta_k, \pi_k) \triangleq \mathbb{E}\left\{ \frac{\partial P_1(y_k[i], y_k[i+1]; \boldsymbol{\theta}_k)}{\partial \boldsymbol{\theta}_k} \frac{\partial P_1(y_k[i], y_k[i+1]; \boldsymbol{\theta}_k)}{\partial \boldsymbol{\theta}_k^T} \right\}$ and the simplification in the first line of Equation (17) arises from exploiting independence of sensor decisions. Finally, with reference to the *BH* scenario, we obtain:

$$\Lambda_{\mathrm{R}} \triangleq \sum_{m=1}^{M} \frac{\partial \ln \left[ P_1(\bar{c}_m[i], \, \bar{c}_m[i+1]; \; \Delta_m, \pi_m) \right]}{\partial \Delta_m} \Bigg|^2_{(\Delta_0, \hat{\pi}_{m,0})} \left[ \mathcal{I}_m^{-1}(\Delta_0, \hat{\pi}_{m,0}) \right]_{\Delta_m, \Delta_m} \tag{18}$$

$$= 2K \sum_{m=1}^{M} \frac{\rho_m \, \left( \hat{\varphi}_m - \hat{\pi}_{m,0} \right)^2}{\hat{\pi}_{m,0}(1 - \hat{\pi}_{m,0})} \tag{19}$$

where $\boldsymbol{\theta}_m \triangleq \left[ \Delta_m \; \pi_m \right]^T$ and $\mathcal{I}_m(\Delta_m, \pi_m) \triangleq \mathbb{E}\left\{ \frac{\partial P_1(\bar{c}_m[i], \bar{c}_m[i+1]; \boldsymbol{\theta}_m)}{\partial \boldsymbol{\theta}_m} \frac{\partial P_1(\bar{c}_m[i], \bar{c}_m[i+1]; \boldsymbol{\theta}_m)}{\partial \boldsymbol{\theta}_m^T} \right\}$.

### 3.4. *Locally Most (Mean) Powerful Test (LM(M)PT)*

Since we are also considering OS testing, a LMPT is appropriate in this context (Kay 1998). Indeed, for H scenario and OS testing, the LMPT is obtained as (Suissa and Shuster 1985).

$$\Lambda_{\mathrm{L}} \triangleq \frac{\partial \ln \left[ P_1(\mathbf{y}[i], \; \mathbf{y}[i+1]; \; \Delta, \pi) \right]}{\partial \Delta} \Bigg|_{(\Delta_0, \hat{\pi}_0)} \sqrt{\left[ \mathcal{I}^{-1}(\Delta_0, \hat{\pi}_0) \right]_{\Delta, \Delta}} = \sqrt{2K} \, \frac{(\hat{\varphi} - \hat{\pi}_0)}{\sqrt{\hat{\pi}_0(1 - \hat{\pi}_0)}} \tag{20}$$

On the other hand, in the *NH* scenario, it can be shown that a LMPT *cannot be obtained* (Kay 1998). Indeed, the first-order Taylor series of the LLR depends on the (unknown) differences $(\Delta_k - \Delta_0) = (\varphi_k - \pi_k)$, which weight the gradient (score) vector $\partial \ln \left[ P_1(\mathbf{y}[i], \; \mathbf{y}[i+1]; \boldsymbol{\Delta}, \boldsymbol{\pi}) \right] / \partial \boldsymbol{\Delta}$. Therefore, to overcome this issue, we resort to a modified multi-dimensional version of the LMPT, which maximizes the mean curvature of the power function in the neighbourhood of $\boldsymbol{\Delta} = \mathbf{0}_K$, that is (Gupta and Vermeire 1986; King and Wu 1997)

$$\Lambda_{\mathrm{L}} \triangleq \sum_{k=1}^{K} \frac{\partial \ln \left[ P_1(y_k[i], \; y_k[i+1]; \; \Delta_k, \pi_k) \right]}{\partial \Delta_k} \Bigg|_{(\Delta_0, \hat{\pi}_{k,0})} \sqrt{\left[ \mathcal{I}_k^{-1}(\Delta_0, \hat{\pi}_{k,0}) \right]_{\Delta_k, \Delta_k}} \tag{21}$$

$$= \sum_{k=1}^{K} \sqrt{2} \, \left\{ y_k[i+1] - y_k[i] \right\}. \tag{22}$$

The statistic in Equation (21) is usually referred[4] to be Locally Most Mean Powerful (LMMP). Remarkably, the LMMP statistic in NH scenario is equivalent to the test based on the statistic $\widehat{\varphi} - \widehat{\pi}_1$, made of the *same ML estimates* as the H scenario.

Finally, in the *BH* scenario, by exploiting problem separability among the regions $\mathcal{R}_m$, $m = 1, \ldots M$, we obtain the following explicit LMMP statistic:

$$\Lambda_{\mathrm{L}} \triangleq \sum_{m=1}^{M} \left. \frac{\partial \ln \left[ P_1(\bar{c}_m[i], \bar{c}_m[i+1];\ \Delta_m, \pi_m) \right]}{\partial \Delta_m} \right|_{(\Delta_0, \hat{\pi}_{m,0})} \sqrt{\left[ \mathcal{I}_m^{-1}(\Delta_0, \hat{\pi}_{m,0}) \right]_{\Delta_m, \Delta_m}} \tag{23}$$

$$= \sqrt{2K} \sum_{m=1}^{M} \sqrt{\rho_m} \frac{\left( \hat{\varphi}_m - \hat{\pi}_{m,0} \right)}{\sqrt{\hat{\pi}_{m,0}(1 - \hat{\pi}_{m,0})}} . \tag{24}$$

### 3.5. Wald Test

The Wald Test is another well-known decision procedure employed in TS testing (Kay 1998). It can be shown that, in H scenario, it equals to (Cressie 1978)

$$\Lambda_{\mathrm{W}} \triangleq \left( \hat{\Delta} - \Delta_0 \right)^2 / \left[ \mathcal{I}^{-1} \left( \hat{\Delta}, \hat{\pi}_1 \right) \right]_{\Delta, \Delta} = K \left( \hat{\varphi} - \hat{\pi}_1 \right)^2 / \left[ \hat{\varphi}(1 - \hat{\varphi}) + \hat{\pi}_1(1 - \hat{\pi}_1) \right] , \tag{25}$$

where $\hat{\Delta} \triangleq (\hat{\varphi} - \hat{\pi}_1)$, i.e. the ML estimate of the probability difference under $\mathcal{H}_1$. The OS counterpart of Wald statistic is instead obtained by replacing $\hat{\Delta}$ with $\hat{\Delta}^+ \triangleq \max\{\hat{\Delta}, 0\}$, thus leading to $\Lambda_{\mathrm{W}} = \left\{ \left[ K \left( \hat{\varphi} - \hat{\pi}_1 \right)^2 \right] / \left[ \hat{\varphi}(1 - \hat{\varphi}) + \hat{\pi}_1(1 - \hat{\pi}_1) \right] \right\} u(\hat{\varphi} - \hat{\pi}_1)$. On the other hand, we remark that in a NH scenario the Wald test *cannot be constructed*. Indeed, in the latter case, the implicit expression (exploiting the mutual independence assumption among the sensors) is:

$$\Lambda_{\mathrm{W}} \triangleq \sum_{k=1}^{K} \left( \hat{\Delta}_k - \Delta_0 \right)^2 / \left[ \mathcal{I}_k^{-1} \left( \hat{\Delta}_k, \hat{\pi}_{k,1} \right) \right]_{\Delta_k, \Delta_k} . \tag{26}$$

Unfortunately, in TS testing the reciprocal of the relevant FIM block $1/ \left[ \mathcal{I}_k^{-1} \left( \hat{\Delta}_k, \hat{\pi}_{k,1} \right) \right]_{\Delta_k, \Delta_k} = 1/ \left[ y_k[i+1](1 - y_k[i+1]) + y_k[i](1 - y_k[i]) \right]$ *always diverges*, thus making the statistic in Equation (26) *not applicable*. A similar reason precludes the application of Wald statistic in NH scenario when testing a OS alternative.

Finally, in the BH scenario, the Wald statistic assumes the explicit expression:

$$\Lambda_{\mathrm{W}} \triangleq \sum_{m=1}^{M} \frac{\left( \hat{\Delta}_m - \Delta_0 \right)^2}{\left[ \mathcal{I}_m^{-1} \left( \hat{\Delta}_m, \hat{\pi}_{m,1} \right) \right]_{\Delta_m, \Delta_m}} = K \sum_{m=1}^{M} \rho_m \frac{\left( \hat{\varphi}_m - \hat{\pi}_{m,1} \right)^2}{\hat{\varphi}_m(1 - \hat{\varphi}_m) + \hat{\pi}_{m,1}(1 - \hat{\pi}_{m,1})} . \tag{27}$$

As for GLR, LMMP and Rao statistics, the result for BH case in Equation (27) merely corresponds to the sum of Wald statistics pertaining to all the regions $\mathcal{R}_m$, $m = 1, \ldots, M$, modelling the surveillance area. Finally, in OS testing, the appropriate Wald statistic is obtained

**Table 1.** Rules comparison: H/BH/NH scenarios and OS/TS testing are considered.

| | GLR | Rao | L(M)MP | Wald |
|---|---|---|---|---|
| **OS/H** | $\mathcal{D}^{\pi}+\mathcal{D}^{\varphi+}$ | $\frac{\chi_0^2}{v(\hat{\pi}_0)}$ | $\frac{\chi_0}{\sqrt{v(\hat{\pi}_0)}}$ | $\frac{\hat{\Delta}^2}{v(\hat{\varphi})+v(\hat{\pi}_1)}u(\hat{\Delta})$ |
| **TS/H** | $\mathcal{D}^{\pi}+\mathcal{D}^{\varphi}$ | $\frac{\chi_0^2}{v(\hat{\pi}_0)}$ | X | $\frac{\hat{\Delta}^2}{v(\hat{\varphi})+v(\hat{\pi}_1)}$ |
| **OS/BH** | $\sum_{m=1}^{M}\mathcal{D}_m^{\pi}+\mathcal{D}_m^{\varphi+}$ | $\sum_{m=1}^{M}\frac{\rho_m\,\chi_{m,0}^2}{v(\hat{\pi}_{m,0})}$ | $\sum_{m=1}^{M}\frac{\sqrt{\rho_m}\,\chi_{m,0}}{\sqrt{v(\hat{\pi}_{m,0})}}$ | $\sum_{m=1}^{M}\frac{\rho_m\,\hat{\Delta}_m^2}{v(\hat{\varphi}_m)+v(\hat{\pi}_{m,1})}u\left(\hat{\Delta}_m\right)$ |
| **TS/BH** | $\sum_{m=1}^{M}\mathcal{D}_m^{\pi}+\mathcal{D}_m^{\varphi}$ | $\sum_{m=1}^{M}\frac{\rho_m\,\chi_{m,0}^2}{v(\hat{\pi}_{m,0})}$ | X | $\sum_{m=1}^{M}\frac{\rho_m\,\hat{\Delta}_m^2}{v(\hat{\varphi}_m)+v(\hat{\pi}_{m,1})}$ |
| **OS/NH** | $\sum_{k=1}^{K}z_{k,i}$ | $\sum_{k=1}^{K}z_{k,i}$ | $\hat{\Delta}$ | X |
| **TS/NH** | $\sum_{k=1}^{K}z_{k,i}$ | $\sum_{k=1}^{K}z_{k,i}$ | X | X |

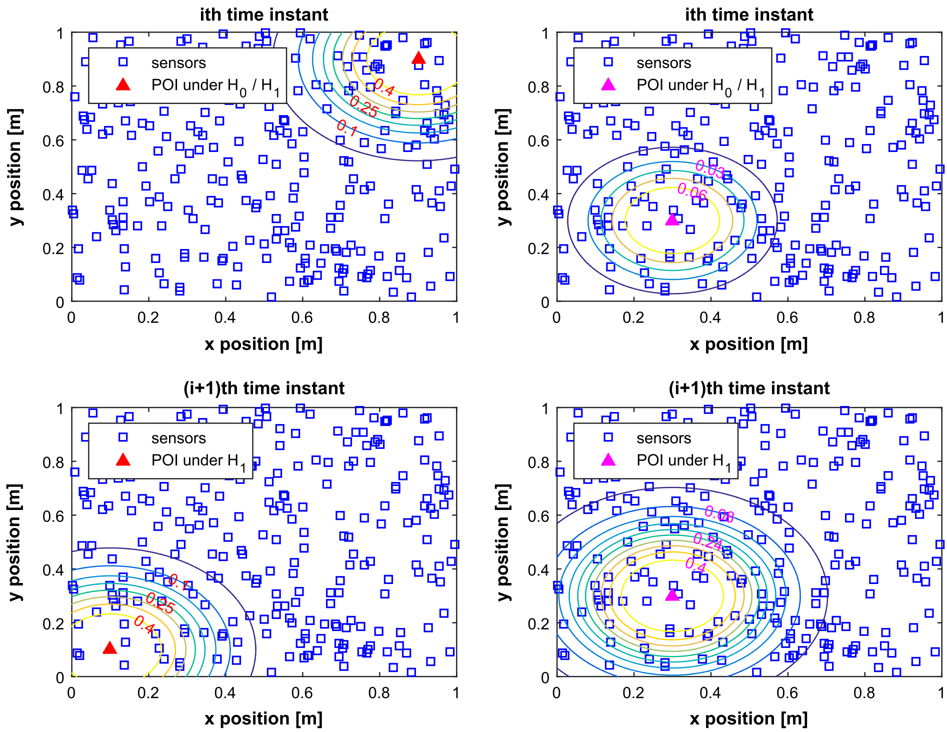
by replacing $\hat{\Delta}_m$ with $\hat{\Delta}_{m,+}\triangleq\max\{\hat{\Delta}_m,0\}$, thus leading to $\Lambda_W = K\sum_{m=1}^{M}\rho_m\{(\hat{\varphi}_m-\hat{\pi}_{m,1})^2/[\hat{\varphi}_m(1-\hat{\varphi}_m)+\hat{\pi}_{m,1}(1-\hat{\pi}_{m,1})]\}\,u(\hat{\varphi}_m-\hat{\pi}_{m,1})$.

## 4. Simulation results

In this section, we investigate the fusion rules within DECHADE framework by focusing on two DCD problems of interest in WSNs: (i) a change of POI position and (ii) an increase of emitted POI power, as detailed in Figure 2.

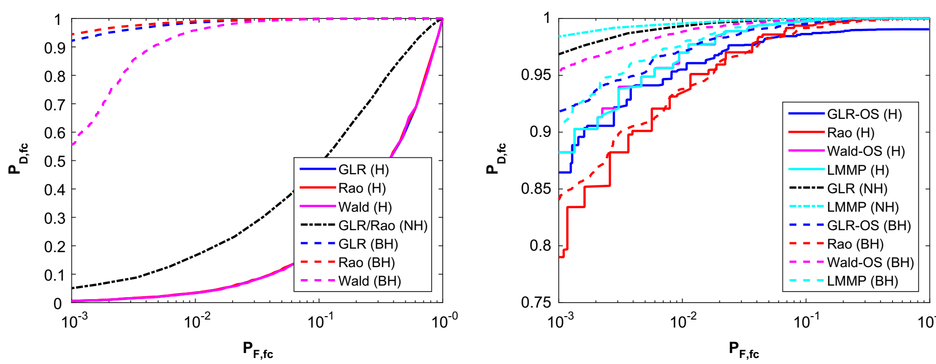
We assume that the sensors monitor a POI within a 2-D surveillance area $\mathcal{A}\triangleq[0,1]^2$ (the nodes are displaced in a random fashion over $\mathcal{A}$, for simplicity) having an isotropic, randomly fluctuating, partially specified spatial signature with distance-dependent path-loss. More specifically, the $k$th sensor measurement at the $\ell$th time instant ($m_k[\ell]\in\mathbb{R}$) adheres to the model $m_k[\ell]=\xi_k[\ell]\,g(\boldsymbol{x}_T[\ell],\boldsymbol{x}_k)+w_k[\ell]$, where the fading coefficient $\xi_k[\ell]\sim\mathcal{N}(0,\theta[\ell])$ models fluctuations in the received signal strength of the (same) POI signature, having an *unknown deterministic* transmitted power $\theta[\ell]$ at $\ell$th instant, which well suits to the case of a realistic POI. Also, $\boldsymbol{x}_T[\ell]\in\mathbb{R}^2$ denotes the *unknown* POI position at the $\ell$th time instant, while $\boldsymbol{x}_k\in\mathbb{R}^2$ denotes the *known* $k$th sensor position, with the pair $(\boldsymbol{x}_T[\ell],\boldsymbol{x}_k)$ *uniquely* determining the value of $g(\boldsymbol{x}_T[\ell],\boldsymbol{x}_k)$, denoting the Amplitude Attenuation Function (AAF); finally, $w_k[\ell]\sim\mathcal{N}(0,\sigma_w^2)$ denotes the sensing noise (without loss of generality, we set here $\sigma_w^2=1$). Hence, in view of these assumptions, $m_k[\ell]\sim\mathcal{N}(0,\,\theta[\ell]\,g^2(\boldsymbol{x}_T[\ell],\boldsymbol{x}_k)+\sigma_w^2)$ holds.

Accordingly, each sensor has been set to reveal the POI based on its local UMP test, corresponding to an *energy test* reporting $y_k[\ell]=1$ (resp. $y_k[\ell]=0$) when $m_k^2[\ell]>\gamma_k$ (resp. $m_k^2[\ell]\leq\gamma_k$) (Guerriero, Svensson, and Willett 2010; Ciuonzo and Salvo Rossi 2017). Consequently, the bit probability of $k$th sensor is $P_k=\Pr\{m_k^2\geq\gamma_k\}=2\,\mathcal{Q}\left(\sqrt{\gamma_k/\{\sigma_w^2+\sigma_s^2\,g^2(\boldsymbol{x}_T,\boldsymbol{x}_k)\}}\right)$ and the threshold $\gamma_k$ is chosen to ensure $2\,\mathcal{Q}\left(\sqrt{\gamma_k/\sigma_w^2}\right)=\bar{P}_{0,k}$ (here set to $\bar{P}_{0,k}=10^{-2}$), so as to constrain the number of sensor false reports when there is no POI to monitor within the area. Here a power-law AAF is chosen $g(\boldsymbol{x}_T,\boldsymbol{x}_k)\triangleq 1\,/\sqrt{1+\left(\|\boldsymbol{x}_T-\boldsymbol{x}_k\|\,/\,\eta\right)^{\alpha}}$, where the POI extent and decay exponent are set as $\eta=0.2$ and $\alpha=4$, respectively. For convenience, we define the POI Signal-To-Noise Ratio (SNR) at $\ell$th time instant as $\mathrm{SNR}[\ell]\triangleq 10\,\log_{10}(\theta[\ell]/\sigma_w^2)$.




(a) Position change setup. (b) Power emission change setup.

**Figure 2.** WSN with $K = 250$ sensors randomly displaced within $\mathcal{A}$. Setup (a): the POI position change is modelled as $\boldsymbol{x}_T[i] = [0.9\ 0.9]^T$, $\boldsymbol{x}_T[i+1] = [0.1\ 0.1]^T$ and $\mathrm{SNR}[i] = \mathrm{SNR}[i+1] = 10\,\mathrm{dB}$. Setup (b): the emitted POI power change is modelled as $\boldsymbol{x}_T[i] = \boldsymbol{x}_T[i+1] = [0.3\ 0.3]^T$, $\mathrm{SNR}[i] = 0\,\mathrm{dB}$ and $\mathrm{SNR}[i+1] = 10\,\mathrm{dB}$.




(a) Position change setup. (b) Power emission change setup.

**Figure 3.** $P_{\mathrm{D,fc}}$ vs. $P_{\mathrm{F,fc}}$ for all the considered rules; for scenario (a) only TS rules (or TS counterparts) of considered rules are considered, whereas for (b) OS rules (or OS-counterparts) have been employed (except for Rao statistic).

Scenario (a) – From inspection of Figure 3(a), it can be concluded that rules designed under BH scenario are the most appealing, with Rao test performing best. This is mainly due to the implicit TS nature of the considered problem (i.e. a change of position may reflect in a either an increase or decrease of the bit/detection probability at each sensor in the WSN). Also, GLR/Rao tests under NH scenario perform satisfactorily, suffering, however, a non-negligible performance loss. Finally, poor performance of H rules in this scenario are due to the fact a change in position does not reflect a global average change of bit probability in the WSN, and this change can be appreciated only if region-based surveillance is performed (i.e. the BH assumption).

Scenario (b) – By looking at Figure 3(b), it can be inferred that all the considered rules within DECHADE framework perform satisfactorily, Rao test in H/BH scenarios performing the worst, as the latter is a rule intrinsically designed for TS testing and therefore does not match with a power increase (which reflects in an increase of bit/detection probability at all the sensors). On the other hand, rules specifically designed for OS testing (or, equivalently, their OS counterparts, e.g. GLR and Wald) reach high performance. The best performance are achieved by the NH group (LMMP and GLR), followed by the BH and H groups (LMMP, GLR and Wald).

## 5. Conclusions and future directions

In this paper, we provided an overview of the classic problem of testing the change in samples drawn from independent Bernoulli PMFs, when the bit probabilities are *not known*, with application to WSNs. Both OS and TS testing were considered, as well as *spatially* H, NH and BH scenarios, in our analysis. Since UMP test was confirmed not to exist in all the cases considered, we derived GLRT, Rao test, Wald test and LM(M)PT while underlining their possible coincidence and/or statistical equivalences, so as to provide a framework of viable rules for DCD, here denoted as DECHADE. Finally, simulation results pertaining to two relevant DCD problems in WSNs were provided, so as to compare the considered rules, underlining the appeal of BH-originated rules under both cases. Asynchronous DCD is a natural extension of the proposed framework and will be considered for future work.

## Notes

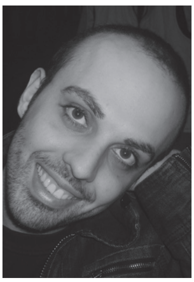
1. *Notation* – Lower-case bold letters denote vectors, with $a_n$ being the $n$th element of $\boldsymbol{a}$; upper-case calligraphic letters, e.g. $\mathcal{A}$, denote finite sets; $\mathbb{E}\{\cdot\}$, $(\cdot)^T$, $\bigvee$ and $\oplus$ denote expectation, transpose, logical OR and XOR, respectively; $P(\cdot)$ denotes probability mass function (PMF), while $P(\cdot|\cdot)$ is the corresponding conditional counterpart; $\mathcal{B}(p)$ denotes a Bernoulli PMF with success probability $p$; $\mathcal{N}(\mu, \sigma^2)$ denotes a Gaussian PDF with mean $\mu$ and variance $\sigma^2$; $\mathcal{Q}(\cdot)$ denotes the complementary CDF of a standard normal random variable, i.e. $\mathcal{N}(0, 1)$; $u(\cdot)$ denotes the Heaviside unit-step function; $\mathcal{D}_{\mathrm{KL}}(\cdot||\cdot)$ denotes the Kullback-Leibler (KL) divergence between distributions (Cover and Thomas 2006). The notation $(\hat{\cdot})$ is denoted to indicate the Maximum Likelihood (ML) estimate of the unknown parameter $(\cdot)$. Also, the symbol $\sim$ means "distributed as" and "$\propto$" is used to underline statistical equivalence between decision statistics. Finally, the notation $\boldsymbol{\delta}_1 \geq .\boldsymbol{\delta}_2$ means that each element of $\boldsymbol{\delta}_1$ is greater or equal than the corresponding element of $\boldsymbol{\delta}_2$, and at least one element of $\boldsymbol{\delta}_1$ is strictly greater than the corresponding element of $\boldsymbol{\delta}_2$.

2. All sensors transmit directly to the FC (which is assumed within their communication range). The framework may be extended to other network topologies (e.g. tandem, hierarchical) with fusion rules to be applied to the aggregation nodes. This aspect is beyond the scope of the paper.
3. The following short-hand notation has been employed: $\mathcal{D}^{\pi} \triangleq \mathcal{D}_{\mathrm{KL}}(\mathcal{B}(\hat{\pi}_1)||\mathcal{B}(\hat{\pi}_0))$, $\mathcal{D}^{\varphi} \triangleq \mathcal{D}_{\mathrm{KL}}(\mathcal{B}(\hat{\varphi})||\mathcal{B}(\hat{\pi}_0))$, $\mathcal{D}^{\varphi+} \triangleq \mathcal{D}_{\mathrm{KL}}(\mathcal{B}(\hat{\varphi}^{+})||\mathcal{B}(\hat{\pi}_0))$, $\mathcal{D}_m^{\pi} \triangleq \mathcal{D}_{\mathrm{KL}}(\mathcal{B}(\hat{\pi}_{m,1})||\mathcal{B}(\hat{\pi}_{m,0}))$, $\mathcal{D}_m^{\varphi} \triangleq \mathcal{D}_{\mathrm{KL}}(\mathcal{B}(\hat{\varphi}_m)||\mathcal{B}(\hat{\pi}_{m,0}))$, $\mathcal{D}_m^{\varphi+} \triangleq \mathcal{D}_{\mathrm{KL}}(\mathcal{B}(\hat{\varphi}_m^{+})||\mathcal{B}(\hat{\pi}_{m,0}))$, $z_{k,i} \triangleq y_k[i] \oplus y_k[i+1]$, $\chi_0 \triangleq (\widehat{\varphi} - \widehat{\pi}_0)$, $\chi_{m,0} \triangleq (\widehat{\varphi}_m - \widehat{\pi}_{m,0})$ and $v(p) \triangleq p(1-p)$.
4. With a slight abuse of notation we will use the symbol $\Lambda_{\mathrm{L}}$ to denote both LMPT and LMMPT, depending on the specific scenario.

## Disclosure statement

No potential conflict of interest was reported by the authors.

## Notes on contributors



***Domenico Ciuonzo*** is currently a researcher at Network Measurement and Monitoring (NM2), Naples, Italy. He holds a PhD in Electronic Engineering from the University of Campania "L. Vanvitelli", Italy and, from 2011, he has held a number of visiting researcher appointments. Since 2014 he is member of the editorial board of different IEEE, IET and Elsevier journals. His research interests fall within the areas of data fusion, statistical signal processing, wireless sensor networks and machine learning. He is a senior member of IEEE.

***Pierluigi Salvo Rossi*** was born in Naples, Italy, on 26 April 1977. He received the Doctor Engineering degree in telecommunications engineering (summa cum laude) and the PhD degree in computer engineering, in 2002 and 2005, respectively, both from the Univ. Naples "Federico II", Italy. From 2005 to 2008, he worked as a postdoc at the Department of Computer Science & Systems, University of Naples "Federico II", Italy, at the Department of Information Engineering, Second University of Naples, Italy, and at the Department of Electronics & Telecommunications, Norwegian University of Science and Technology (NTNU), Trondheim, Norway. From 2008 to 2014, he was an assistant professor (tenured in 2011) in telecommunications at the Department of Industrial & Information Engineering, Second University of Naples, Italy. From 2014 to 2016, he was an associate professor in signal processing with the Department of Electronics & Telecommunications, NTNU, Norway. From 2016 to 2017, he was a full professor in signal processing with the Department of Electronic Systems, NTNU, Norway. Since 2017, he is a principal engineer with the Advanced Analytics & Machine Learning Team, Kongsberg Digital AS, Norway. He held visiting appointments at the Department of Electrical & Computer Engineering, Drexel University, Philadelphia, PA, US, at the Department of Electrical & Information Technology, Lund University, Sweden, at the Department of Electronics & Telecommunications, NTNU, Norway, and at the Excellence Center for Wireless Sensor Networks, Uppsala University, Sweden. He is an IEEE senior member and serves as a senior editor for the IEEE Communication Letter (since 2016) and an associate editor for the IEEE Transactions and Wireless Communication (since 2015). He was an associate editor for the IEEE Communication Letter (from 2012 to 2016) and a Guest Editor for Elsevier Physical Communication. (2012). His research interests fall within the areas of communications and signal processing.

## References

Banerjee, T., and V. V. Veeravalli. 2015. "Data-efficient Quickest Change Detection in Sensor Networks." *IEEE Transactions on Signal Processing* 63 (14): 3727–3735.

Basin, M., P. Shi, and D. Calderon-Alvarez. 2010. "Central Suboptimal $H_\infty$ Filter Design for Linear Time-varying Systems with State and Measurement Delays." *International Journal of Systems Science* 41 (4): 411–421.

Caballero-Águila, R., A. Hermoso-Carazo, and J. Linares-Pérez. 2015. "Optimal State Estimation for Networked Systems with Random Parameter Matrices, Correlated Noises and Delayed Measurements." *International Journal of General Systems* 44 (2): 142–154.

Ciuonzo, D., A. De Maio, and P. Salvo Rossi. 2015. "A Systematic Framework for Composite Hypothesis Testing of Independent Bernoulli Trials." *IEEE Signal Processing Letters* 22 (9): 1249–1253.

Ciuonzo, D., G. Papa, G. Romano, P. Salvo Rossi, and P. Willett. 2013. "One-bit Decentralized Detection with a Rao Test for Multisensor Fusion." *IEEE Signal Processing Letters* 20 (9): 861–864.

Ciuonzo, D., and P. Salvo Rossi. 2014. "Decision Fusion with Unknown Sensor Detection Probability." *IEEE Signal Processing Letters* 21 (2): 208–212.

Ciuonzo, D., and P. Salvo Rossi. 2017. "Distributed Detection of a Non-cooperative Target via Generalized Locally-optimum Approaches." *Information Fusion* 36: 261–274.

Ciuonzo, D., P. Salvo Rossi, and P. Willett. 2017. "Generalized Rao Test for Decentralized Detection of an Uncooperative Target." *IEEE Signal Processing Letters* 24 (5): 678–682.

Cover, T. M., and J. A. Thomas. 2006. *Elements of Information Theory*. New York: Wiley-Interscience.

Cressie, N. 1978. "Testing for the Equality of Two Binomial Proportions." *Annals of the Institute of Statistical Mathematics* 30 (1): 421–427.

Faivishevsky, L. 2016. "Information Theoretic Multivariate Change Detection for Multisensory Information Processing in Internet of Things." In *IEEE ICASSP*, Shangai, China, 6250–6254.

Guerriero, M., L. Svensson, and P. Willett. 2010. "Bayesian Data Fusion for Distributed Target Detection in Sensor Networks." *IEEE Transactions on Signal Processing* 58 (6): 3417–3421.

Gupta, A. S., and L. Vermeire. 1986. "Locally Optimal Tests for Multiparameter Hypotheses." *Journal of the American Statistical Association* 81 (395): 819–825.

He, T., S. Ben-David, and L. Tong. 2006. "Nonparametric Change Detection and Estimation in Large-scale Sensor Networks." *IEEE Transactions on Signal Processing* 54 (4): 1204–1217.

Hoeffding, W. 1965. "Asymptotically Optimal Tests for Multinomial Distributions." *The Annals of Mathematical Statistics* 36 (2): 369–401.

Kay, S. M. 1998. *Fundamentals of Statistical Signal Processing, Volume 2: Detection Theory*. Englewood Cliffs, NJ: Prentice Hall PTR.

King, M. L., and P. X. Wu. 1997. "Locally Optimal One-sided Tests for Multiparameter Hypotheses." *Econometric Reviews* 16 (2): 131–156.

Niu, R., and P. K. Varshney. 2008. "Performance Analysis of Distributed Detection in a Random Sensor Field." *IEEE Transactions on Signal Processing* 56 (1): 339–349.

Patwari, N., A. O. Hero, and B. M. Sadler. 2003. "Hierarchical Censoring Sensors for Change Detection." In *IEEE SSP Workshop*, St. Louis, MO, USA, 21–24.

Suissa, S., and J. S. Shuster. 1985. "Exact Unconditional Sample Sizes for the $2 \times 2$ Binomial Trial." *Journal of the Royal Statistical Society Series A (General)* 148 (4): 317–327.

Varshney, P. K. 1996. *Distributed Detection and Data Fusion*. 1st ed. Springer-Verlag New York Inc.