Time-Aware Distributed Sequential Detection of Gas Dispersion via Wireless Sensor Networks

Gianluca Tabella, *Graduate Student Member, IEEE*, Domenico Ciuonzo, *Senior Member, IEEE*, Yasin Yilmaz, *Senior Member, IEEE*, Xiaodong Wang, *Fellow, IEEE*, and Pierluigi Salvo Rossi, *Senior Member, IEEE*

***Abstract*—This work addresses the problem of detecting gas dispersions through concentration sensors with wireless transmission capabilities organized as a distributed Wireless Sensor Network (WSN). The concentration sensors in the WSN perform local sequential detection (SD) and transmit their individual decisions to the Fusion Center (FC) according to a transmission rule designed to meet the low-energy requirements of a wireless setup. The FC receives the transmissions sent by the sensors and makes a more reliable global decision by employing a SD algorithm. Two variants of the SD algorithm named *Continuous Sampling Algorithm* (CSA) and *Decision-Triggered Sampling Algorithm* (DTSA), each with its own transmission rule, are presented and compared against a fully-batch algorithm named *Batch Sampling Algorithm* (BSA). The CSA operates as a *time-aware* detector by incorporating the time of each transmission in the detection rule. The proposed framework encompasses the gas dispersion model into the FC's decision rule and leverages real-time weather measurements. The case study involves an accidental dispersion of carbon dioxide (CO_2). System performances are evaluated in terms of the receiver operating characteristic (ROC) curve as well as average decision delay and communication cost.**

***Index Terms*—Wireless sensor networks (WSN), sequential detection (SD), distributed detection, Industry 4.0, gas dispersion.**

I. Introduction

WIRELESS Sensor Networks (WSNs) have become increasingly popular for monitoring applications in the past decade: a trend that was amplified with the emergence of the Internet of Things (IoT) paradigm [3]. One area of interest has been the detection of *harmful events*, with applications related to i) security, counter-terrorism, and defense [4], and ii) safety and environmental protection in Industry 4.0 [5], [6].

More specifically, WSNs are typically composed of low-cost devices monitoring the surrounding environment. Due to stringent bandwidth and/or energy constraints (e.g. to ensure the long-lasting lifetime of IoT nodes), sensors are usually required to send extremely-compressed versions of their measurements to a Fusion Center (FC) which collects and analyzes the data for a final decision. For this reason, the detection of diffusive sources in safety-critical systems via WSNs has shifted toward the adoption of binary sensors [7], [8]. In such scenarios, the FC generates an alarm if an adverse event is detected, triggering appropriate actions to mitigate the consequences. This is particularly relevant to manufacturing, energy, and process industries, where equipment malfunctions can put workers and the environment in danger, as well as result in unplanned shutdowns, high costs, and lost revenue [9].

In this context, the associated inference problems involve the *early* detection of uncooperative sources, such as the loss of containment of fluids in the process industry (in gas and/or liquid form). The detection of *heavy gases* is among the most relevant problems, as heavy gases *do not adhere* to neutral or positively-buoyant dispersion behavior and tend to spread along the ground, with the further threat of asphyxiation induced by the displacement of air, resulting in low oxygen concentrations. In these industrial scenarios, it is of utmost importance to accurately detect such critical events as quickly as possible. An additional source of complexity must be taken into account in case the gas of interest is commonly found in the atmosphere: this can sensibly decrease the detector's performance. To this end, an *industrial IoT* setup with inexpensive sensors and the possibility of exploiting real-time weather data as well as the integration of the gas dispersion model represents an enabler for this problem.

This work addresses the *sequential detection* (SD) of gas dispersion using a network of wireless concentration sensors, focusing on gases with a non-null atmospheric concentration in normal conditions. Performance evaluation is carried out on a simulated dispersion of heavy gas. More specifically, in this study, we adopt the SD framework with the aim of achieving higher accuracy and lower detection time with respect to a fully

Manuscript received 17 April 2023; revised 7 September 2023; accepted 8 October 2023. Date of publication 13 October 2023; date of current version 27 October 2023. Part of this work has been presented to the 2022 IEEE 12th Sensor Array and Multichannel Signal Processing Workshop (SAM) [DOI: 10.1109/SAM53842.2022.9827715] and the 2022 25th International Conference on Information Fusion (FUSION) [DOI: 10.23919/FUSION49751.2022.9841386]. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Yan Chen. *(Corresponding author: Gianluca Tabella.)*

Gianluca Tabella is with the Department of Electronic Systems, Norwegian University of Science and Technology, 7034 Trondheim, Norway (e-mail: gianluca.tabella@ntnu.no).

Domenico Ciuonzo is with the Department of Electrical Engineering and Information Technologies, University of Naples "Federico II," 80138 Naples, Italy (e-mail: domenico.ciuonzo@unina.it).

Yasin Yilmaz is with the Department of Electrical Engineering, University of South Florida, Tampa, FL 33620 USA (e-mail: yasiny@usf.edu).

Xiaodong Wang is with the Department of Electrical Engineering, Columbia University, New York, NY 10027 USA (e-mail: wangx@ee.columbia.edu).

Pierluigi Salvo Rossi is with the Department of Electronic Systems, Norwegian University of Science and Technology, 7034 Trondheim, Norway, and also with the Department of Gas Technology, SINTEF Energy Research, 7034 Trondheim, Norway (e-mail: salvorossi@ieee.org).

Digital Object Identifier 10.1109/TSIPN.2023.3324586

batch approach. In SD the observations are processed one at a time, and a decision is made after each observation to either declare the presence or absence of the event of interest or continue with the detection process.

A. Related Work

Several methods for gas detection have been developed assuming a Gaussian-plume point-source model based on diffusion/advection processes (not suitable for heavy gases) or direct use of Fick's laws of diffusion (not suitable for complex systems), e.g. with application to dispersion of biochemical moving sources [10], [11], atmospheric pollutants [12] and release of light gases [13]. Also, in order to deal with the vague prior, importance sampling was implemented using the progressive correction technique in [14]. The algorithm showed good performance in terms of both localization and estimation accuracy. An interesting feature of this approach is that system-level performance can be controlled by a local detection threshold. Other novel methods rely on neural networks for plume tracking [15]. However, such works neglect the detection task and directly focus on the characterization of the dispersion which is facilitated by the use of a centralized sensor network.

Nevertheless, the current literature lacks studies on the use of WSNs for the detection of gases with a non-null atmospheric concentration in normal conditions, e.g. carbon dioxide (CO_2) whose current average concentration in the atmosphere is around 400 ppm. Most of the above-mentioned studies focus on the detection of gases that are not commonly present in the atmosphere, leading to amplified signal-to-noise ratios. Some preliminary investigations have attempted to fill this gap by demonstrating that the use of a model-based algorithm implemented through a WSN can improve performances in contrast to a model-free algorithm (i.e. the implementation of a *counting rule* on the received binary decisions) [1], [2]. This study builds upon these initial inquiries by incorporating the issue of *early* detection, achieved through the implementation of a SD approach.

Event detection can be tackled with multiple approaches. Distributed detection via WSNs using *batch* decision rules is a mature area of research [16], [17], [18], [19], [20]. SD (also known as *sequential analysis* or *sequential hypothesis testing*) is a well known framework popularized by Wald with the *sequential probability ratio test* (SPRT) [21], [22]. The optimality of SPRT allows achieving faster online decisions with respect to traditional *batch* detectors requiring a fixed sample size before decisions can be made via the *likelihood ratio test* (LRT) [23]. A complete overview of SD can be found in [21], [23], [24].

SD via WSNs has been explored in the last decade, but still remains an open research topic. In [25], an architecture was proposed where both the sensors and the FC perform sequential detection with sensors communicating their respective local decisions to the FC. Such a setup was proven to have asymptotically equivalent performance to the centralized counterpart in specific conditions. A higher-performance alternative was presented later in [26], grounded on the assumption that the observed signal is a sampled version of a continuous stochastic process with continuous paths. Other works (e.g. [27]) applied the distributed SD paradigm to develop spectrum sensing schemes for cognitive radio networks exploring quantization strategies. Practical aspects such as imperfect reporting channels (between sensors and FC) and requirements for reduced energy consumption were considered in [28], [29]. Recent works have focused on alternative tests than the SPRT to be used in WSNs, as the exact knowledge of the distribution function of the signal in the alternative hypothesis is often missing. Therefore, for a composite hypothesis test suitable in WSNs, a *generalized sequential probability ratio test* (GSPRT) was studied in [30].

Truncated versions of sequential tests have been explored in order to bind the decision time that might otherwise become undesirably long. When applied to one-sided tests, they are usually referred to as *truncated one-sided* (TOS) tests. A solid overview of truncated tests can be found in [24]. This option was firstly explored for SPRT and GSPRT in [31], and recently adopted in combinations with other tests. More specifically, truncation was applied to the *repeated significance test* in [32], to the *random distortion test* in [33], and finally to a FC performing the *score test* in the context of detection of a non-cooperative moving target in [34], [35].

B. Contribution and Paper Organization

This work investigates the use of a WSN made of concentration sensors in an industrial IoT setup with inexpensive small-battery sensors for gas detection purposes. First, we introduce a fully-batch algorithm, named *Batch Sampling Algorithm* (BSA), characterized by a fixed sample size at both sensors and FC. Next, with the goal of reducing the detection time, we propose two fully sequential algorithms. In the proposed strategies, each sensor measures the local concentration and takes a local decision via SD regarding the presence or absence of a gas dispersion. A *transmission rule* is present to regulate the communication from the sensors to the FC. Next, the FC, based on the transmissions received by the sensors, performs a global decision taking advantage of updated weather measurements and the integration of the gas dispersion model in the detection rule.

The first proposed method, named *Decision-Triggered Sampling Algorithm* (DTSA), has the FC sampling the sensors' transmission only when local decisions are taken. The second proposed method, named *Continuous Sampling Algorithm* (CSA), requires the FC to continuously monitor the transmissions from the sensors (which also encodes the temporary lack of a local decision). In the CSA, at each instant, the FC updates a test statistic based on the transmission values and the time elapsed since the last sensors' decision, resulting in a *time-aware* algorithm.

This work presents *new advances* in the field of industrial monitoring as listed in the following:

- The study is based on the integration of the gas dispersion model into the design of the FC;
- The proposed methods make use of externally-available measurements from weather stations (e.g. wind measurements);
- The sequential nature of the proposed methods allows to reduce the detection time and removes the limitation

imposed by a fixed number of samples needed to take a decision;
- The introduction of a transmission rule tailored for sequential binary local detectors allows for reduced energy consumption in the case of CSA.

This work further explores the use of WSNs for gas detection via the integration of the dispersion model within the detection algorithm previously presented in [1], [2]. In these earlier works, we compared the well-known model-free *counting fusion rule* with a model-aware *generalized Chair-Varshney fusion rule*, proving the benefits of such implementation. The further contribution given by this work is the extension from a single-sample detection to a SD approach.

The remainder of the article is organized as follows: Section II provides a system overview, focusing on the WSN architecture and the signal model (including the gas dispersion characterizations); the batch approach is described in Section III; the proposed sequential algorithms are described in Section IV, focusing on both sensors and the FC; Section V discusses the performances of the local sequential detectors in terms of accuracy, and decision delay; the computational complexity and the communication costs are discussed in Section VI; numerical results of the considered case study are presented in Section VII; finally, conclusions and further works are addressed in Section VIII.

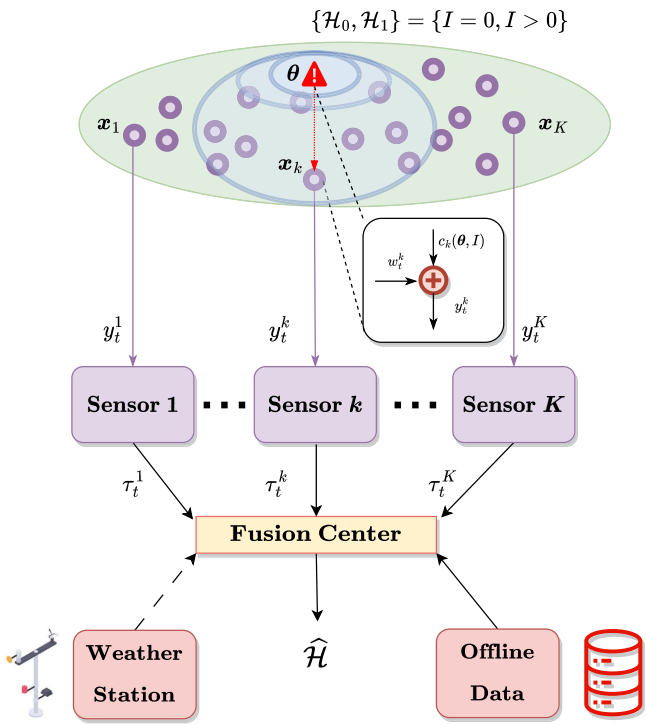


Fig. 1. Wireless sensor network architecture.

C. Notation

Uppercase (resp. lowercase) bold letters denote matrices (resp. column vectors); $[\cdot]^{\mathrm{T}}$ denotes the transpose operator; $\hat{a}$ is an estimate of the variable a; $\mathbb{E}(\cdot)$, $\mathbb{V}\mathrm{ar}(\cdot)$, $\mathbb{C}\mathrm{ov}(\cdot\,,\cdot)$ denote expectation, variance, and covariance; $\mathbb{P}(\cdot)$ and $p(\cdot)$ denote probability mass functions (PMFs) and probability density functions (PDFs), while $\mathbb{P}(\cdot|\cdot)$ and $p(\cdot|\cdot)$ their corresponding conditional counterparts; in particular, $\mathbb{E}_j(\cdot)$, $\mathbb{P}_j(\cdot)$ and $p_j(\cdot)$ denote the expectation, the PMF, and PDF, respectively, under the hypothesis $\mathcal{H}_j$, with $j \in \{0, 1\}$; $\mathcal{L}_\theta(a) \triangleq \ln[\mathbb{P}_1(a;\theta)/\mathbb{P}_0(a)]$ is the log-likelihood ratio where the dependence on the parameter θ is highlighted; $\mathcal{U}(a, b)$ denotes a continuous uniform distribution with minimum value a and maximum value b; $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes a multivariate Gaussian distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$; $\mathcal{Q}(\cdot)$ is the complementary cumulative distribution function (CCDF) of the standard normal distribution; $\delta_{a,b}$ is the Kronecker delta; finally $\mathcal{O}(\cdot)$ denotes the big O notation.

II. System Model

What follows is the overview of the distributed WSN under consideration, followed by the characterization of the signal measured by the sensors.

A. Wireless Sensor Network Model

The scenario consists of a distributed WSN comprising K static sensors and its task is to assess the *global* absence ($\mathcal{H}_0$) or presence ($\mathcal{H}_1$) of a gas leak within the monitored environment (a schematic representation is given in Fig. 1).[1] Such a dispersion is characterized by its position $\boldsymbol{\theta}$ and intensity I. For the kth sensor ($k = 1, \ldots, K$), the location and the measurement of gas concentration at discrete-time $t \in \mathbb{N}^+$ are denoted by $\boldsymbol{x}_k$ and y_t^k, respectively. Each sensor computes a test statistic on the above-mentioned signal and assesses the *local* absence ($\mathcal{H}_0$) or presence ($\mathcal{H}_1$) of an anomalous excessive gas concentration. For the sake of convenience, we assumed the sensor to have the same sampling frequency and to be perfectly synchronized. In the algorithms under study, when a sensor makes a decision, it immediately starts a new detection instance until the FC takes a global decision, allowing the FC to receive multiple decisions from a single sensor. The global decision exploits the integration of real-time weather data as well as the dispersion model of the gas.

When the BSA is employed, each sensor takes a decision after a fixed number of measurements. At each instant, each sensor sends a *transmission value* $\tau_t^k = 1$ (resp. $\tau_t^k = -1$) to the FC if $\mathcal{H}_1$ is declared (resp. $\mathcal{H}_0$), or $\tau_t^k = 0$ if the sensor has not finished collecting its fixed number of samples. Specifically, when $\tau_t^k = 0$, the sensor *does not* transmit a physical communication to the FC. At a predetermined time, the FC takes a global decision $\widehat{\mathcal{H}} \in \{\mathcal{H}_0, \mathcal{H}_1\}$ computing a test statistic on the received values $\{\tau_t^k : |\tau_t^k| = 1\}_{k,t}$.

In the newly proposed methods (DTSA and CSA), both the sensors and the FC make use of SD, with the aim of reducing the decision delay obtained in the BSA.

In the DTSA, the sensors send a transmission value to the FC after completing a sequential test, i.e. not at predetermined

[1]The possibility of incurring faulty sensors is not taken into account as it is outside of the scope of the present work. Fault detection and identification techniques based on data-driven philosophy could be readily incorporated in the proposed approach [36].

times, unlike in the BSA. Here, at each t, the FC performs a test on $\{\tau_t^k : |\tau_t^k| = 1\}_k$ and takes a global decision.

In the CSA, each individual sensor transmits a bit $\tau_t^k = 1$ (resp. $\tau_t^k = 0$) to the FC if $\mathcal{H}_1$ is locally declared (resp. if $\mathcal{H}_0$ is locally declared or if the sensor has not reached a decision yet). In addition to being spectrally efficient, as only one bit is transmitted on the communication channel between the sensor and the FC, such a system is highly energy efficient when OOK modulation is employed for communicating the local decisions [5]. Moreover, at each t, the FC sequentially performs a *time-aware* test on the transmission values $\{\tau_t^k, a_t^k\}_k$, where a_t^k is the number of instants passed since the last decision made by the kth sensor.

In this work, we assume a perfect communication channel between sensors and FC.

B. Signal Model

The statistical model of the measured gas concentration y_t^k, depending on the corresponding hypothesis, is the following:

$$\begin{cases} \mathcal{H}_0: & y_t^k = w_t^k \\ \mathcal{H}_1: & y_t^k = c_k + w_t^k \end{cases}, \tag{1}$$

where $w_t^k \sim \mathcal{N}(\mu_k, \sigma_k^2)$ represents the gas concentration present in normal operating conditions in the surrounding of the kth sensor [37], where the values of μ_k and σ_k^2 are both *known*. The values of $\{\mu_k\}_k$ and $\{\sigma_k^2\}_k$ can be estimated by calculating the sample mean and sample variance from a set of measurements acquired in normal operating conditions ($\mathcal{H}_0$). Also, $c_k \geq 0$ is the observed excess gas concentration resulting from dispersion, here assumed constant in time since this work deals with steady-state dispersions.

In this work, we assume that the measurements collected by the same sensor $\{y_t^k\}_t$ are i.i.d., while the measurements collected by different sensors $\{y_t^k\}_k$ are independent, with distributions that vary depending on $\{c_k\}_k$. This assumption arises from (1) and the treatment of $\{w_t^k\}_{t,k}$ as i.i.d. variables. Although this treatment simplifies reality, assuming null space and time correlation in the modeling of $\{w_t^k\}_{t,k}$ can be justified by ensuring adequate spatial separation between the sensors and a sufficiently low sampling frequency. A low sampling frequency results in auto-covariance values dominated by lower-frequency components. Moreover, accurately predicting these lower-frequency components in the atmospheric fluctuation of the concentration of the gas presents significant complexities. Therefore, we chose to simplify the model by excluding them [20], [38]. Hence, the distribution of y_t^k is:

$$\begin{cases} \mathcal{H}_0: & y_t^k \sim \mathcal{N}\left(\mu_k, \sigma_k^2\right) \\ \mathcal{H}_1: & y_t^k \sim \mathcal{N}\left(\mu_k + c_k, \sigma_k^2\right) \end{cases}, \tag{2}$$

where the value of c_k is the result of a dispersion phenomenon. There is extensive literature on how to obtain the value of c_k due to its industrial safety applications. We assume:

$$c_k = \mathcal{F}(\boldsymbol{x}_k, \mathcal{A}, \mathcal{B}, \mathcal{C}), \tag{3}$$

where $\mathcal{A}$ is the set of all *unknown* variables such as the release position ($\boldsymbol{\theta}$), and the intensity (I); $\mathcal{B}$ is the set of variables whose value is *known* and constant in time, once the variables in $\mathcal{A}$ are fixed. $\mathcal{B}$ includes variables such as temperature, density, initial concentration of the release, as well as morphological properties of the area. $\mathcal{C}$ is the set of variables that can be considered independent from the variables in $\mathcal{A}$ and $\boldsymbol{x}_k$, and whose value is *known* via real-time measurement. This set includes the meteorological parameters. The values of the variables in $\mathcal{B}$ are set by exploiting the knowledge of the monitored environment, while those in $\mathcal{C}$ require real-time meteorological data. For the case of a release, the most important variables belonging to $\mathcal{A}$ are $\boldsymbol{\theta}$ and I, hence once $\boldsymbol{x}_k$, $\boldsymbol{\theta}$, and I are fixed, and the variables in $\mathcal{B}$ and $\mathcal{C}$ are available, the value of c_k can be unequivocally determined.

III. Batch Detection

This section is dedicated to the BSA which relies on fixed sample size at both sensors and FC levels for the detection task. This algorithm is designed for a FC that is able to compute c_k via the map in (3) once the unknown dispersion variables belonging to $\mathcal{A}$ have been fixed. c_k is written as $c_k(\theta, I)$ to emphasize the unknown variables. The other variables in (3) are known and constant throughout the detection procedure. Real-time weather data and, possibly, physical knowledge of the monitored area are necessary to determine the variables in $\mathcal{B}$ and $\mathcal{C}$. The considered architecture requires solving a maximization problem: we assume grid-search optimization.

Specifically, the kth sensor will take a local decision after collecting $\mathcal{T}_k$ samples, after which it restarts a new detection instance. As a consequence, each sensor is characterized by a deterministic *stopping time* corresponding to the mth decision $t_m^k = m\mathcal{T}_k$. For the model described in (1), the *likelihood ratio test* (LRT) is *uniformly most powerful*, resulting in each sensor calculating a statistic Λ_t^k with the following form:[2]

$$\Lambda_t^k \triangleq \sum_{i=1}^{t} \left(y_i^k - \mu_k\right) = \begin{cases} \lambda_1^k, & t = 1 \\ \Lambda_{(t-1)}^k + \lambda_t^k, & t > 1 \end{cases}, \tag{4}$$

where $\lambda_t^k \triangleq y_t^k - \mu_k$. This leads to the following decision rule:

$$d_m^k \triangleq \begin{cases} \mathcal{H}_1, & \text{if } \Lambda_{m\mathcal{T}_k}^k - \Lambda_{(m-1)\mathcal{T}_k}^k \geq \gamma_k \\ \mathcal{H}_0, & \text{otherwise} \end{cases}, \tag{5}$$

with γ_k as a local test threshold. The *probability of false alarm* ($\mathcal{P}_F^k$) and *detection* ($\mathcal{P}_D^k(c_k)$) of the local batch detector are the following:

$$\begin{aligned} \mathcal{P}_F^k &\triangleq \mathbb{P}_0\left(d_m^k = \mathcal{H}_1\right) = \mathbb{P}_0\left(d_1^k = \mathcal{H}_1\right) \\ &= \mathbb{P}_0\left(\Lambda_{\mathcal{T}_k}^k \geq \gamma_k\right) = \mathcal{Q}\left(\frac{\gamma_k}{\sqrt{\mathcal{T}_k \sigma_k^2}}\right), \\ \mathcal{P}_D^k(c_k) &\triangleq \mathbb{P}_1\left(d_m^k = \mathcal{H}_1\right) = \mathcal{Q}\left(\frac{\gamma_k - \mathcal{T}_k c_k}{\sqrt{\mathcal{T}_k \sigma_k^2}}\right), \end{aligned} \tag{6}$$

where we exploited the fact that $\{\Lambda_{m\mathcal{T}_k}^k - \Lambda_{(m-1)\mathcal{T}_k}^k\}_m$ are i.i.d., and therefore we chose $m = 1$.

[2]The LRT statistic can be simplified into $\sum_{i=1}^{t} y_i^k$. However, we prefer using the above-mentioned statistic to ease the comparison with the DTSA and CSA.

To prove that the above-mentioned elements are i.i.d., it is sufficient to say that, since $\{y_i^k\}_i$ are i.i.d., the same applies to the disjoint subsequences $\{y^k_{(m-1)\mathcal{T}_k+1}, \ldots, y^k_{m\mathcal{T}_k}\}_m$. This means that the set of statistics computed over these subsequences, expressed as $\{\sum_{i=(m-1)\mathcal{T}_k+1}^{m\mathcal{T}_k} \lambda_i^k\}_m$, also consist of i.i.d. elements. Finally, by applying the definition of Λ_t^k from (4), we can conclude that the elements in the sequence $\{\Lambda^k_{m\mathcal{T}_k} - \Lambda^k_{(m-1)\mathcal{T}_k}\}_m$ are i.i.d. as well.

At each instant t, the kth sensor sends a transmission value to the FC according to the following transmission rule:

$$\tau_t^k \triangleq \begin{cases} +1\,, & \text{if } \exists\, m : t = m\,\mathcal{T}_k \ \wedge\ d_m^k = \mathcal{H}_1 \\ -1\,, & \text{if } \exists\, m : t = m\,\mathcal{T}_k \ \wedge\ d_m^k = \mathcal{H}_0 \\ 0\,, & \text{otherwise} \end{cases}, \tag{7}$$

with $\tau_t^k = 0$ indicating the absence of physical communication from the sensor to FC during the sensor's fixed decision period.

At predetermined time $\mathcal{T}^*$, the FC takes a global decision employing a GLRT statistic Λ^{B} on the local decisions:

$$\widehat{\mathcal{H}} \triangleq \begin{cases} \mathcal{H}_1\,, & \text{if } \Lambda^{\mathrm{B}} \geq \gamma^* \\ \mathcal{H}_0\,, & \text{otherwise} \end{cases}, \tag{8}$$

with γ^* as a global test threshold. Since the statistic is a function of the local decisions having known fixed decision delays, it is recommended to set $\mathcal{T}^*$ so that $\exists\, k, a \in \mathbb{N}^+ : \mathcal{T}^* = a\,\mathcal{T}_k$. In order to calculate Λ^{B}, the FC has to keep track of the transmission times t_m^k:

$$t_m^k = \inf\left\{t > t_{m-1}^k : \left|\tau_t^k\right| = 1\right\}, \quad t_0^k = 0\,. \tag{9}$$

Specifically, Λ^{B} is a statistic on the received local decisions, which translates into a test on the transmission values $\{\tau_t^k : |\tau_t^k| = 1\}_{t,k}$:

$$\begin{aligned} \Lambda^{\mathrm{B}} &\triangleq \max_{\boldsymbol{\theta}, I} \left\{ \mathcal{L}_{\boldsymbol{\theta},I}\left(\left\{\tau_t^k : \left|\tau_t^k\right| = 1\right\}_{\substack{1\leq t\leq \mathcal{T}^* \\ 1\leq k\leq K}} \right) \right\} \\ &= \max_{\boldsymbol{\theta}, I} \left\{ \sum_{k=1}^{K} \sum_{m=1}^{\mathcal{M}_{\mathcal{T}^*}^k} \mathcal{L}_{\boldsymbol{\theta},I}\left(\tau_{t_m^k}^k\right) \right\}, \end{aligned} \tag{10}$$

where $\mathcal{M}_{\mathcal{T}^*}^k \triangleq \sum_{t=1}^{\mathcal{T}^*} |\tau_t^k|$ is the number of local decisions taken by the kth sensor up to time $\mathcal{T}^*$. In particular, $\mathcal{L}_{\boldsymbol{\theta},I}(\tau_{t_m^k}^k)$ has the following form:

$$\mathcal{L}_{\boldsymbol{\theta},I}\left(\tau_{t_m^k}^k\right) = \begin{cases} \ln \frac{\mathcal{P}_D^k(c_k(\theta,I))}{\mathcal{P}_F^k}\,, & \text{if } \tau_{t_m^k}^k = +1 \\ \ln \frac{1-\mathcal{P}_D^k(c_k(\theta,I))}{1-\mathcal{P}_F^k}\,, & \text{if } \tau_{t_m^k}^k = -1 \end{cases}. \tag{11}$$

IV. Sequential Detection

In this section, we explore the CSA and DTSA. First, we examine the SD algorithm at a sensor level shared by both architectures. Next, we outline the algorithm at the FC level in the two different methods.

A. Local Sequential Detection

Each sensor performs SD on the hypotheses in (1). Equation (2) highlights that the test has to be *one-sided* since $\{\mathcal{H}_0, \mathcal{H}_1\}$ correspond to $\{c_k = 0, c_k \geq 0\}$, respectively. For this task, we compute the GSPRT statistic, where the parameter c_k in the log-likelihood ratio is replaced with its maximum likelihood estimate $\widehat{c}_{k,t} \triangleq \frac{1}{t}\sum_{i=1}^t y_i^k - \mu_k$. This results in the same statistic Λ_t^k already introduced in (4).

The GSPRT, analogously to the *generalized likelihood ratio test*, is asymptotically non-negative for one-sided hypothesis testing problems, thus the use of a negative threshold is unfeasible. To overcome this issue, we resort to a TOS test by establishing the maximum amount of time $\mathcal{T}_k$ between two consecutive local decisions that the kth sensor can take in order to declare $\mathcal{H}_1$, otherwise, $\mathcal{H}_0$ is declared. Denoting γ_k as a positive local threshold and the time at which the sensor takes the mth decision with t_m^k, the mth stopping time is defined as the following:

$$\begin{aligned} t_m^k &\triangleq \min\left\{\inf\left\{t > t_{m-1}^k : \Lambda_t^k - \Lambda_{t_{m-1}^k}^k \geq \gamma_k\right\}, t_{m-1}^k + \mathcal{T}_k\right\} \\ &= \min\left\{\inf\left\{t > t_{m-1}^k : \sum_{i=t_{m-1}^k+1}^{t} \lambda_i^k \geq \gamma_k\right\}, t_{m-1}^k + \mathcal{T}_k\right\}, \end{aligned} \tag{12}$$

with $t_0^k = 0$ and $\Lambda_0^k = 0$. Next, the decision rule is as follows:

$$d_m^k \triangleq \begin{cases} \mathcal{H}_1\,, & \text{if } \Lambda_{t_m^k}^k - \Lambda_{t_{m-1}^k}^k \geq \gamma_k \\ \mathcal{H}_0\,, & \text{otherwise} \end{cases}. \tag{13}$$

Remarks: In the process of deriving the local detector, we employ the Karlin-Rubin theorem to reduce the test statistic (via monotonic transformations) before substituting c_k with its MLE. This reduction is achieved by exploiting the non-negative nature of c_k.

B. Fusion Center Sequential Detection

Here we describe the two FC detection methods for gas dispersion: i) DTSA, a SD algorithm with the FC performing a test statistic solely based on the received local decisions (the knowledge of the sampling period is not required); ii) CSA, a novel *time-aware* SD algorithm with the FC performing a test statistic on those instants where the sensors take decisions as well as on those instants where the sensors have not reached a decision yet (the knowledge of the sampling period for each sensor is required). As in the BSA, both methods rely on the ability to calculate the values of c_k via the map in (3).

1) Decision-Triggered Sampling Algorithm (DTSA): This algorithm consists of the FC sequentially updating a test statistic when a local decision is taken. Similarly to the BSA, the transmission rule encodes the detection status of the sensors:

$$\tau_t^k \triangleq \begin{cases} +1\,, & \text{if } \exists\, m : t = t_m^k \ \wedge\ d_m^k = \mathcal{H}_1 \\ -1\,, & \text{if } \exists\, m : t = t_m^k \ \wedge\ d_m^k = \mathcal{H}_0 \\ 0\,, & \text{otherwise} \end{cases}, \tag{14}$$

where $\tau_t^k = 0$ indicates the absence of a physical transmission from the sensor to FC. This transmission rule translates into a test statistic performed on those transmission values that are decision-triggered $\{\tau_t^k : |\tau_t^k| = 1\}_{t,k}$. Similarly to the design of the local detectors, the presence of the unknown parameters $\boldsymbol{\theta}$ and I in the hypothesis $\mathcal{H}_1$ requires the use of a GSPRT statistic

here denoted by Λ_t^{D}. Likewise, we use a time limit $\mathcal{T}^*$ at which, if the FC has not declared $\mathcal{H}_1$ yet, $\mathcal{H}_0$ is automatically reported, leading to the following stopping rule and decision rule:

$$t^* \triangleq \min\left\{\inf\left\{t : \Lambda_t^{\mathrm{D}} \geq \gamma^*\right\}, \mathcal{T}^*\right\}, \tag{15}$$

$$\widehat{\mathcal{H}} \triangleq \begin{cases} \mathcal{H}_1\,, & \text{if } \Lambda_{t^*}^{\mathrm{D}} \geq \gamma^* \\ \mathcal{H}_0\,, & \text{otherwise} \end{cases}. \tag{16}$$

At each t, in order to calculate Λ_t^{D}, the FC needs to recover the stopping times t_m^k (using (9)), as well as (recursively) calculate the number of local decisions taken by the kth sensor up to time t, for all k:

$$\mathcal{M}_t^k = \mathcal{M}_{t-1}^k + \left|\tau_t^k\right|, \quad \mathcal{M}_0^k = 0\,. \tag{17}$$

The next step consists of the FC computing the GSPRT statistic Λ_t^{D}:

$$\begin{aligned}\Lambda_t^{\mathrm{D}} &\triangleq \max_{\boldsymbol{\theta},I}\left\{\mathcal{L}_{\boldsymbol{\theta},I}\left(\left\{\tau_i^k : \left|\tau_i^k\right| = 1\right\}_{\substack{1\leq i\leq t\\ 1\leq k\leq K}}\right)\right\} \\ &= \max_{\boldsymbol{\theta},I}\left\{\sum_{k=1}^{K}\sum_{m=1}^{\mathcal{M}_t^k}\mathcal{L}_{\boldsymbol{\theta},I}\left(\tau_{t_m^k}^k\right)\right\}, \end{aligned} \tag{18}$$

where again we exploited the independence of the local decisions in time and space. The term $\mathcal{L}_{\boldsymbol{\theta},I}(\tau_{t_m^k}^k)$ can be obtained using the *overall local performances* of the sensors:

$$\mathcal{L}_{\boldsymbol{\theta},I}\left(\tau_{t_m^k}^k\right) = \begin{cases} \ln\frac{\mathcal{P}_D^k(c_k(\theta,I))}{\mathcal{P}_F^k}\,, & \text{if } \tau_{t_m^k}^k = +1 \\ \ln\frac{\mathcal{P}_M^k(c_k(\theta,I))}{\mathcal{P}_C^k}\,, & \text{if } \tau_{t_m^k}^k = -1 \end{cases}, \tag{19}$$

with $\mathcal{P}_D^k$, $\mathcal{P}_F^k$, $\mathcal{P}_M^k$, and $\mathcal{P}_C^k$ representing the *overall probability of detection*, *false alarm*, *miss detection*, and *correct rejection*, respectively, of a sequential detector. These metrics are discussed in Section V.

2) Continuous Sampling Algorithm (CSA): In this configuration, the kth sensor transmits a message to the FC only when $\mathcal{H}_1$ is declared, so we can state the following transmission rule at each t:

$$\tau_t^k \triangleq \begin{cases} 1\,, & \text{if } \exists\, m : t = t_m^k \ \wedge\ d_m^k = \mathcal{H}_1 \\ 0\,, & \text{otherwise} \end{cases}, \tag{20}$$

where τ_t^k is the *transmission value*, with $\tau_t^k = 0$ indicating the absence of a physical transmission from the sensor to FC.

Meanwhile, the FC sequentially updates a statistic using the received transmission values $\{\tau_t^k\}_{k,t}$. The knowledge of the sampling period of each sensor allows such a continuous sampling although $\tau_t^k = 0$ does not constitute a physical transmission. The reason behind the use of the same transmission value $\tau_t^k = 0$ to represent the absence of a decision and a negative decision lies in the deterministic nature of the time taken by a sensor to declare $\mathcal{H}_0$ (equal to $\mathcal{T}_k$) which allows to unequivocally distinguish the two cases. Similarly to the previously proposed architecture, we employ a truncated GSPRT, whose statistic is indicated with Λ_t^{C} with a time limit $\mathcal{T}^*$, leading to the following stopping rule and decision rule:

$$t^* \triangleq \min\left\{\inf\left\{t : \Lambda_t^{\mathrm{C}} \geq \gamma^*\right\}, \mathcal{T}^*\right\},$$

$$\widehat{\mathcal{H}} \triangleq \begin{cases} \mathcal{H}_1\,, & \text{if } \Lambda_{t^*}^{\mathrm{C}} \geq \gamma^* \\ \mathcal{H}_0\,, & \text{otherwise} \end{cases}. \tag{21}$$

At each t, the calculation of Λ_t^{C} requires the FC to sequentially deduce, for each sensor, whether the received transmission value τ_t^k corresponds to a local decision or not, and retrieve the current delay a_t^k:

$$t_m^k = \min\left\{\inf\left\{t > t_{m-1}^k : \tau_t^k = 1\right\}, t_{m-1}^k + \mathcal{T}_k\right\}, \tag{22}$$

$$\mathcal{M}_t^k = \begin{cases} \mathcal{M}_{t-1}^k + 1\,, & \text{if } t = t_{\mathcal{M}_{t-1}^k+1}^k \\ \mathcal{M}_{t-1}^k\,, & \text{otherwise} \end{cases}, \tag{23}$$

$$a_t^k = \begin{cases} 1\,, & \text{if } t = t_{\mathcal{M}_{t-1}^k+1}^k \\ a_{t-1}^k + 1\,, & \text{otherwise} \end{cases}, \tag{24}$$

where $\mathcal{M}_t^k$ now counts the number of local decisions taken by the kth sensor at time t *including* the one that is currently being taken, with $t_0^k = 0$ and $\mathcal{M}_0^k = 0$. The next step consists of the FC computing the GSPRT statistic Λ_t^{C}:

$$\begin{aligned}\Lambda_t^{\mathrm{C}} &\triangleq \max_{\boldsymbol{\theta},I}\left\{\mathcal{L}_{\boldsymbol{\theta},I}\left(\left\{\tau_i^k\right\}_{\substack{1\leq i\leq t\\ 1\leq k\leq K}}\right)\right\} \\ &= \max_{\boldsymbol{\theta},I}\left\{\sum_{k=1}^{K}\sum_{m=1}^{\mathcal{M}_t^k}\mathcal{L}_{\boldsymbol{\theta},I}\left(\tau_{\min\{t,t_m^k\}}^k, a_{\min\{t,t_m^k\}}^k\right)\right\}, \end{aligned} \tag{25}$$

where we exploited the independence of the local decisions in time and space. The generic value of $\mathcal{L}_{\boldsymbol{\theta},I}(\tau_t^k, a_t^k)$ can be expressed using the *instant local performances* of the sensors:

$$\mathcal{L}_{\boldsymbol{\theta},I}\left(\tau_t^k, a_t^k\right) = \begin{cases} \ln\frac{\mathcal{P}_D^{(k,a_t^k)}(c_k(\theta,I))}{\mathcal{P}_F^{(k,a_t^k)}}\,, & \text{if } \tau_t^k = 1 \\ \ln\frac{\mathcal{P}_M^{(k,a_t^k)}(c_k(\theta,I))}{\mathcal{P}_C^{(k,a_t^k)}}\,, & \text{if } \tau_t^k = 0 \end{cases}, \tag{26}$$

with $\mathcal{P}_D^{(k,i)}$, $\mathcal{P}_F^{(k,i)}$, $\mathcal{P}_M^{(k,i)}$, and $\mathcal{P}_C^{(k,i)}$ representing the *instant probability of detection*, *false alarm*, *miss detection*, and *correct rejection*, respectively (see Section V).

V. Analysis of Local Sequential Detection

The assessment of the local performances of sequential detectors is now reported. First, we assess the *instant* and *overall performances*. Next, we analyze the local decision delays. In the rest of the work, to ease the comparison between the presented architectures, we will assume that the deadlines $\mathcal{T}_k$'s (resp. $\mathcal{T}^*$) used in the DTSA and CSA are set to have the same values of the sample sizes at sensor level (resp. FC level) used in the BSA.

A. Instant Local Performances

We analyze the local performances at the kth sensor in terms of *instant probability of false alarm* $\mathcal{P}_F^{(k,i)}$ and *instant probability of detection* $\mathcal{P}_D^{(k,i)}(c_k)$ for the generic mth decision with respect to each time instant $\{t_{m-1}^k + i\}_{i=1}^{\mathcal{T}_k}$:

$$\begin{aligned}\mathcal{P}_F^{(k,i)} &\triangleq \mathbb{P}_0\left(d_m^k = \mathcal{H}_1, t_m^k - t_{m-1}^k = i\right) \\ &= \mathbb{P}_0\left(\left\{\Lambda_t^k < \gamma_k\right\}_{t<i}, \Lambda_i^k \geq \gamma_k\right), \end{aligned}$$

$$\mathcal{P}_D^{(k,i)}(c_k) \triangleq \mathbb{P}_1\Big(\big\{\Lambda_t^k < \gamma_k\big\}_{t<i}, \Lambda_i^k \geq \gamma_k; c_k\Big)\,, \tag{27}$$

where, analogously to (6), we exploited the fact that $\{\Lambda_t^k - \Lambda_{t_{m-1}^k}^k\}_m$ are i.i.d. for any $t \in [t_{m-1}+1, t_{m-1}+i]$, and therefore we chose $m = 1$.

By examining (2) and (4), we have:

$$\begin{cases} \mathcal{H}_0: & \Lambda_i^k \sim \mathcal{N}\big(0, i\sigma_k^2\big) \\ \mathcal{H}_1: & \Lambda_i^k \sim \mathcal{N}\big(ic_k, i\sigma_k^2\big) \end{cases}. \tag{28}$$

Therefore (27) is obtained via computing the CDF of a multivariate Gaussian random variable in the form of $\mathbb{P}_j(\boldsymbol{z}_i^k \leq \mathbf{0})$, with $\boldsymbol{z}_i^k \overset{\mathcal{H}_j}{\sim} \mathcal{N}(\boldsymbol{\mu}_{\boldsymbol{z}_i^k}^j, \boldsymbol{\Sigma}_{\boldsymbol{z}_i^k})$, where:

$$\boldsymbol{z}_i^k \triangleq \begin{bmatrix} \Lambda_1^k - \gamma_k \\ \vdots \\ \Lambda_{i-1}^k - \gamma_k \\ -\Lambda_i^k + \gamma_k \end{bmatrix}, \boldsymbol{\mu}_{\boldsymbol{z}_i^k}^0 \triangleq \begin{bmatrix} -\gamma_k \\ \vdots \\ -\gamma_k \\ \gamma_k \end{bmatrix},$$

$$\boldsymbol{\mu}_{\boldsymbol{z}_i^k}^1 \triangleq \begin{bmatrix} c_k - \gamma_k \\ \vdots \\ (i-1)c_k - \gamma_k \\ -ic_k + \gamma_k \end{bmatrix},$$

$$\boldsymbol{\Sigma}_{\boldsymbol{z}_i^k} \triangleq \begin{bmatrix} \sigma_k^2 & \sigma_k^2 & \cdots & \sigma_k^2 & -\sigma_k^2 \\ \sigma_k^2 & 2\sigma_k^2 & \cdots & 2\sigma_k^2 & -2\sigma_k^2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \sigma_k^2 & 2\sigma_k^2 & \cdots & (i-1)\sigma_k^2 & -(i-1)\sigma_k^2 \\ -\sigma_k^2 & -2\sigma_k^2 & \cdots & -(i-1)\sigma_k^2 & i\sigma_k^2 \end{bmatrix}. \tag{29}$$

Moreover, computing the *instant probability of correct rejection* ($\mathcal{P}_C^{(k,i)}$) and the *instant probability of miss detection* ($\mathcal{P}_M^{(k,i)}(c_k)$) is needed for the CSA:

$$\mathcal{P}_C^{(k,i)} \triangleq 1 - \sum_{j=1}^{i} \mathcal{P}_F^{(k,j)}\,, \quad \mathcal{P}_M^{(k,i)}(c_k) \triangleq 1 - \sum_{j=1}^{i} \mathcal{P}_D^{(k,j)}(c_k)\,. \tag{30}$$

Hence, the values of $\mathcal{P}_C^{(k,i)}$ and $\mathcal{P}_M^{(k,i)}$ are obtained from previously calculated probabilities. However, such probabilities are computed via numerical methods (being CDFs of multivariate Gaussian random variables). Unless the approximation error is sufficiently low, we might experience (mainly for high values of i) an accumulation of errors in the final result, especially undesirable when leading to negative values in (30). For this reason, we also include the direct calculation of $\mathcal{P}_C^{(k,i)}$ and $\mathcal{P}_M^{(k,i)}(c_k)$ which are defined as:

$$\mathcal{P}_C^{(k,i)} \triangleq \mathbb{P}_0\Big(\big\{\Lambda_t^k < \gamma_k\big\}_{t\leq i}\Big)\,,$$
$$\mathcal{P}_M^{(k,i)}(c_k) \triangleq \mathbb{P}_1\Big(\big\{\Lambda_t^k < \gamma_k\big\}_{t\leq i}; c_k\Big)\,. \tag{31}$$

These can be obtained computing $\mathbb{P}_j(\boldsymbol{v}_i^k \leq \mathbf{0})$, with $\boldsymbol{v}_i^k \overset{\mathcal{H}_j}{\sim} \mathcal{N}(\boldsymbol{\mu}_{\boldsymbol{v}_i^k}^j, \boldsymbol{\Sigma}_{\boldsymbol{v}_i^k})$. Specifically:

$$\boldsymbol{v}_i^k \triangleq \begin{bmatrix} \Lambda_1^k - \gamma_k \\ \vdots \\ \Lambda_i^k - \gamma_k \end{bmatrix}, \boldsymbol{\mu}_{\boldsymbol{v}_i^k}^0 \triangleq \begin{bmatrix} -\gamma_k \\ \vdots \\ -\gamma_k \end{bmatrix}, \boldsymbol{\mu}_{\boldsymbol{v}_i^k}^1 \triangleq \begin{bmatrix} c_k - \gamma_k \\ \vdots \\ (i-1)c_k - \gamma_k \\ ic_k - \gamma_k \end{bmatrix},$$

$$\boldsymbol{\Sigma}_{\boldsymbol{v}_i^k} \triangleq \begin{bmatrix} \sigma_k^2 & \sigma_k^2 & \sigma_k^2 & \cdots & \sigma_k^2 \\ \sigma_k^2 & 2\sigma_k^2 & 2\sigma_k^2 & \cdots & 2\sigma_k^2 \\ \sigma_k^2 & 2\sigma_k^2 & 3\sigma_k^2 & \cdots & 3\sigma_k^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_k^2 & 2\sigma_k^2 & 3\sigma_k^2 & \cdots & i\sigma_k^2 \end{bmatrix}. \tag{32}$$

$\boldsymbol{\Sigma}_{\boldsymbol{z}_i^k}$ and $\boldsymbol{\Sigma}_{\boldsymbol{v}_i^k}$ are derived in Appendix A.

B. Overall Local Performances

The *overall probabilities of false alarm* ($\mathcal{P}_F^k$) and *detection* ($\mathcal{P}_D^k(c_k)$) at the kth sensor for the mth decision are:

$$\mathcal{P}_F^k \triangleq \mathbb{P}_0\big(d_m^k = \mathcal{H}_1\big) = \sum_{i=1}^{\mathcal{T}_k} \mathcal{P}_F^{(k,i)}\,,$$

$$\mathcal{P}_D^k(c_k) \triangleq \mathbb{P}_1\big(d_m^k = \mathcal{H}_1; c_k\big) = \sum_{i=1}^{\mathcal{T}_k} \mathcal{P}_D^{(k,i)}(c_k)\,. \tag{33}$$

The results in (30) do not relate to local decisions except for $i = \mathcal{T}_k$, in such case, the overall probabilities of correct rejection ($\mathcal{P}_C^k$) and miss detection ($\mathcal{P}_M^k(c_k)$) are readily given:

$$\mathcal{P}_C^k \triangleq \mathcal{P}_C^{(k,\mathcal{T}_k)}\,, \quad \mathcal{P}_M^k(c_k) \triangleq \mathcal{P}_M^{(k,\mathcal{T}_k)}(c_k)\,. \tag{34}$$

C. Local Decision Delays

With the local detection algorithm being sequential, one can evaluate the average time taken to reach a decision. We use $\mathcal{D}_{1j}^k$ to represent the expected time taken by the kth sensor to declare $\mathcal{H}_1$ when $\mathcal{H}_j$ is true, while $\mathcal{D}_{0X}^k$ refers to the declaration of $\mathcal{H}_0$ independently of the true hypothesis:

$$\mathcal{D}_{10}^k \triangleq \mathbb{E}_0\big(t_m^k - t_{m-1}^k \big| d_m^k = \mathcal{H}_1\big)$$
$$= \sum_{i=1}^{\mathcal{T}_k} i\,\mathbb{P}_0\big(t_1^k = i \big| d_1^k = \mathcal{H}_1\big) = \frac{1}{\mathcal{P}_F^k} \sum_{i=1}^{\mathcal{T}_k} i\,\mathcal{P}_F^{(k,i)}\,,$$

$$\mathcal{D}_{11}^k(c_k) \triangleq \frac{1}{\mathcal{P}_D^k(c_k)} \sum_{i=1}^{\mathcal{T}_k} i\,\mathcal{P}_D^{(k,i)}(c_k)\,,$$

$$\mathcal{D}_{0X}^k \triangleq \mathbb{E}\big(t_m^k - t_{m-1}^k \big| d_m^k = \mathcal{H}_0\big) = \mathcal{T}_k\,. \tag{35}$$

In particular, given $d_m^k = \mathcal{H}_0$, then $\mathcal{D}_{0X}^k = \mathcal{T}_k$ *almost surely*. Moreover, it is possible to express the expected time $\mathcal{D}_{Xj}^k$ taken by the kth sensor to take any decision when $\mathcal{H}_j$ is true:

$$\mathcal{D}_{X0}^k \triangleq \mathbb{E}_0\big(t_m^k - t_{m-1}^k\big) = \sum_{i=1}^{\mathcal{T}_k} i\,\mathbb{P}_0\big(t_1^k = i\big)$$

$$= \mathcal{D}_{10}^k \mathcal{P}_F^k + \mathcal{T}_k \mathcal{P}_C^k = \mathcal{T}_k - \sum_{i=1}^{\mathcal{T}_k} (\mathcal{T}_k - i) \mathcal{P}_F^{(k,i)} \,,$$

$$\mathcal{D}_{X1}^k(c_k) \triangleq \mathcal{D}_{11}^k(c_k) \mathcal{P}_D^k(c_k) + \mathcal{T}_k \mathcal{P}_M^k(c_k)$$

$$= \mathcal{T}_k - \sum_{i=1}^{\mathcal{T}_k} (\mathcal{T}_k - i) \mathcal{P}_D^{(k,i)}(c_k) \,. \tag{36}$$

These expressions explicitly show that the local decision delay of the kth sensor, in the case of sequential detection, is always upper-bounded by $\mathcal{T}_k$.

VI. Computational Complexity

This section assesses the computational complexity and the communication costs associated with the online FC detection algorithm in the case of the BSA, DTSA, and CSA.

A. Offline Preparation

In the three algorithms, the main task of the FC is to sum the observations' log-likelihood ratios and find the maximum with respect to the unknown parameters. Since the online computation of log-likelihood ratios becomes computationally intensive as the grids get *finer*, we assume an offline stage preceding the online detection where the log-likelihood ratios are precomputed for each grid point and uploaded to the FC. The variables in $\mathcal{C}$ are known (thanks to the use of real-time weather data) but vary with time, thus a grid of possible values for those variables is required as well.

Such an offline data preparation has the benefit of reducing the real-time computational toll on the FC, but it suffers from the mismatch between the measured meteorological data and its closest value on the grid. However, such a difference can be arbitrarily reduced using a finer grid for the variables in $\mathcal{C}$ during the offline data computation.

B. Computation Complexity

The three different algorithms contain instructions for the FC on when and how the decision statistic must be updated. Given an instant where the FC is required to update the statistic, we have the same computational complexity $\mathcal{O}(K \cdot |\text{grid}(\boldsymbol{\theta})| \cdot |\text{grid}(I)|)$ across the three algorithms. The main computational difference lies in the rate at which these updates must be carried out which varies according to the employed algorithm. Let us assume that our network consists of a single sensor ($K = 1$): the BSA calculates the statistic only once after $\mathcal{T}^*$ instants (because of its batch nature); the DTSA, instead, has a mean update period of $\mathcal{D}_{X1}^1(c_k)$ when $\mathcal{H}_1$ is true (resp. $\mathcal{D}_{X0}^1$ when $\mathcal{H}_0$ is true) with $\mathcal{T}_1$ as an upper bound (see (36)), while the CSA has an update period equal to 1. We conclude by saying that: $1 \leq \mathcal{D}_{X1}^1 \leq \mathcal{D}_{X0}^1 \leq \mathcal{T}^*$, which shows the higher rate of update of the CSA, followed by the DTSA, both bounded by the BSA. Variations of the CSA might be proposed where the update of the statistic is carried out with a period higher than 1 and lower or equal than $\mathcal{T}_1$ as long as no local positive decision is taken (if the update period is equal to $\mathcal{T}_1$ the update frequency would collide with that of the DTSA). These observations can be extended to networks having $K > 1$.

C. Communication Costs

Each architecture is configured with a distinct combination of decision rule and transmission rule at the sensor level, resulting in a different *average transmission period* (ATP) between physical communications from *each sensor* to the FC. The subsequent results show the average transmission periods for each of the shown architectures:

$$\text{ATP}_{\text{CSA}} \triangleq \mathbb{E}\left(t_b^k - t_a^k \middle| \sum_{t=t_a^k}^{t_b^k} \tau_t^k = 2 \right)$$

$$= \begin{cases} \mathcal{D}_{X0}^k / \mathcal{P}_F^k \,, & \text{if } \mathcal{H}_0 \text{ is true} \\ \mathcal{D}_{X1}^k(c_k) / \mathcal{P}_D^k(c_k) \,, & \text{if } \mathcal{H}_1 \text{ is true} \end{cases} , \tag{37}$$

$$\text{ATP}_{\text{DTSA}} \triangleq \mathbb{E}\left(t_m^k - t_{m-1}^k\right)$$

$$= \begin{cases} \mathcal{D}_{X0}^k \,, & \text{if } \mathcal{H}_0 \text{ is true} \\ \mathcal{D}_{X1}^k(c_k) \,, & \text{if } \mathcal{H}_1 \text{ is true} \end{cases} , \tag{38}$$

$$\text{ATP}_{\text{BSA}} \triangleq t_m^k - t_{m-1}^k = \mathcal{T}_k \,, \quad \textit{almost surely}. \tag{39}$$

The derivation of ATP_{CSA} is reported in Appendix B.

We can immediately observe that $\text{ATP}_{\text{CSA}} \geq \text{ATP}_{\text{DTSA}}$. This is a direct consequence of the absence of physical communication when a sensor decides $\mathcal{H}_0$ in the CSA architecture. We can further notice, using (36), that $\text{ATP}_{\text{BSA}} \geq \text{ATP}_{\text{DTSA}}$. A comparison between ATP_{CSA} and ATP_{BSA} is less trivial and will be discussed via the case study in Section VII.

VII. Results

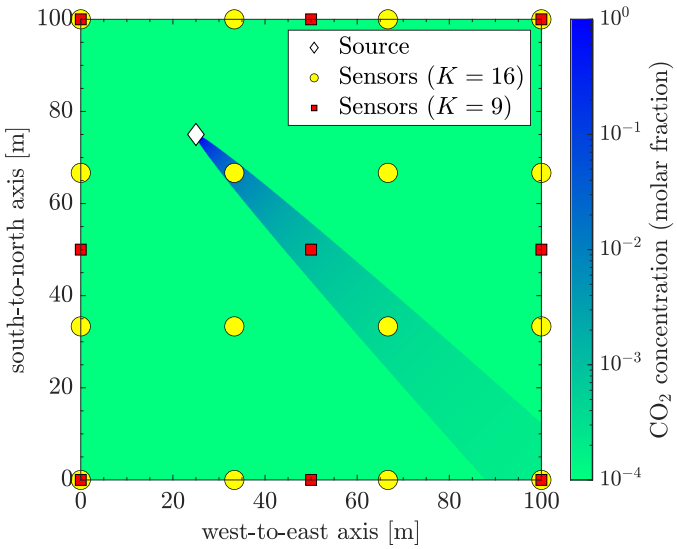
The considered scenario simulates the dispersion of *saturated* carbon dioxide (CO_2), a heavy gas whose density, at atmospheric temperature and pressure, is about 1.5 times larger than the air density. Heavy gases need specialized models that can predict their behavior like the well-known *Britter & McQuaid* (B&M) model for continuous releases [39], [40], [41]. The B&M model is based on the manual reading of a chart which prevents its use by the FC, thus a set of analytical relationships described in [2] is employed to convert it into a set of equations. The B&M's output, with respect to the kth sensor, is c_k. The variables belonging to the sets $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$, for the B&M model are the following:

$$\mathcal{A} = \{\boldsymbol{\theta}, I, D\}, \; \mathcal{B} = \{T, \rho, c_0\}, \; \mathcal{C} = \{T_{\text{atm}}, \rho_{\text{air}}, u, \varphi\}, \tag{40}$$

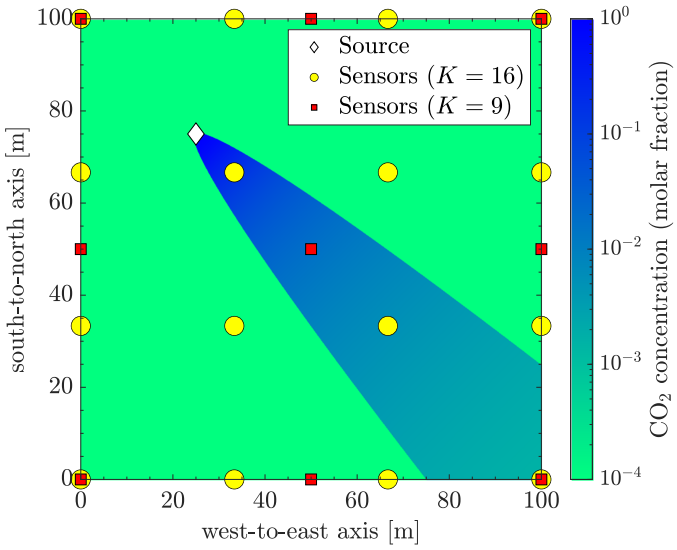
where T, ρ, c_0, and I are the temperature, density, concentration, and intensity of the gas at release condition; D is the release diameter; T_{atm} is the atmospheric temperature; ρ_{air} is the density of air at T_{atm}; finally u and φ are the wind speed at the height of 10 meters[3] and its direction.[4] D is a parameter that, like $\boldsymbol{\theta}$ and I, should be estimated as it is unknown. However, its contribution to the value of c_k is negligible for small values of D (which is the case for accidental dispersion), allowing us to assume it as

[3] If wind speed is available at a different height, several conversion methods are available [42].

[4] Wind blowing from north: 0° (360°), east: 90°, south: 180°, west: 270°.



(a) Low intensity: $I = 0.05\,\mathrm{m}^3/\mathrm{s}$ (maximum value)



(b) High intensity: $I = 0.5\,\mathrm{m}^3/\mathrm{s}$ (maximum value)

Fig. 2. (Mean) concentration maps in a dispersion scenario at different intensities with $\boldsymbol{\theta} = [25\,\mathrm{m}\ 75\,\mathrm{m}]^{\mathrm{T}}$, $\varphi = 315°$, $D = 0.1\,\mathrm{m}$, and $u = 5\,\mathrm{m/s}$.

known and equal to zero reducing the computational complexity of the algorithms.

Here, we assume that both the dispersion model and the signal model in Section II are accurate so that possible differences between the assumptions and the actual phenomenon can be neglected. The evaluation of the consequences of a non-negligible mismatch is outside the scope of this work.

The results are obtained via simulation of a monitored square area with sides of 100 meters with equally-spaced sensors, as shown in Fig. 2. The simulated settings refer to combinations of network size $K \in \{9, 16\}$ each with *low intensity* and *high intensity* dispersions. The corresponding four combinations allow an exhaustive comparison of the proposed algorithms. The results of each combination have been computed via numerical simulation with 10^5 Monte Carlo runs equally divided between

TABLE I
PARAMETERS USED FOR THE SIMULATIONS

Parameter	Value / Distribution	Notes
θ_1 and θ_2	$\mathcal{U}(0, 100)\,\mathrm{m}$	uniform in monitored area
c_0	1	molar vapor fraction ($1 = 10^6$ ppm)
$T^{(\mathrm{op})}$	253 K	[43]
$P^{(\mathrm{op})}$	19.8 bar	saturation pressure at $T^{(\mathrm{op})}$
T	219 K	*Soave-Redlich–Kwong* EOS [44]
P_{atm}	1.01 bar	–
ρ	$2.48\,\mathrm{kg/m}^3$	*Soave-Redlich–Kwong* EOS [44]
T_{atm}	293 K	–
ρ_{air}	$1.20\,\mathrm{kg/m}^3$	–
u	$\mathcal{U}(0, 10)\,\mathrm{m/s}$	–
φ	$\mathcal{U}(0, 2\pi)$	–
I	$\mathcal{U}(0, 0.05)\,\mathrm{m}^3/\mathrm{s}$	low intensity dispersion
I	$\mathcal{U}(0.4, 0.5)\,\mathrm{m}^3/\mathrm{s}$	high intensity dispersion
D	$\mathcal{U}(0, 0.2)\,\mathrm{m}$	–
μ_k	400 ppm	$\forall k$
σ_k	200 ppm	$\forall k$
$\mathcal{T}_k$	4	$\forall k$
$\mathcal{P}_F^k$	0.05	$\forall k$
γ_k	693 ppm	$\forall k$, DTSA and CSA
γ_k	658 ppm	$\forall k$, BSA

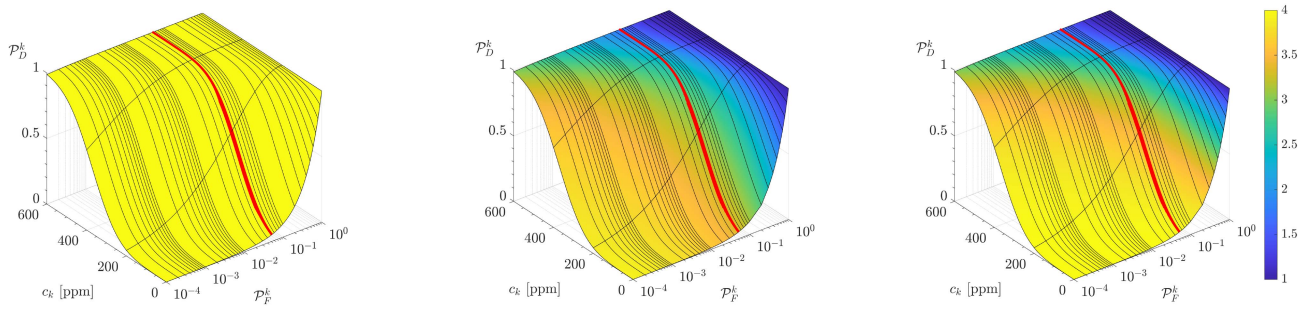
TABLE II
PARAMETERS USED FOR GRID CONSTRUCTION

Parameter	Grid Limits	Grid Interval
θ_1 and θ_2	$[0, 100]\,\mathrm{m}$	$1\,\mathrm{m}$
I (low intensity)	$[0, 0.05]\,\mathrm{m}^3/\mathrm{s}$	$1/60\,\mathrm{m}^3/\mathrm{s}$
I (medium intensity)	$(0.05, 0.4)\,\mathrm{m}^3/\mathrm{s}$	$60/7\,\mathrm{m}^3/\mathrm{s}$
I (high intensity)	$[0.4, 0.5]\,\mathrm{m}^3/\mathrm{s}$	$1/30\,\mathrm{m}^3/\mathrm{s}$
u	$[0.5, 10]\,\mathrm{m/s}$	$0.5\,\mathrm{m/s}$
φ	$[0, 2\pi)$	$\pi/8$

$\mathcal{H}_0$ and $\mathcal{H}_1$ via MATLAB software. At each run, parameters such as wind direction (φ), wind speed (u), dispersion position ($\boldsymbol{\theta}$), intensity (I), and dispersion diameter (D) are generated according to a uniform distribution in a predetermined realistic range of values. The remaining parameters are kept constant across all the runs. The values or the distribution boundaries of the parameters are shown in Table I, while the specifications of the parameter grids necessary for the offline preparation of data are reported in Table II. In the present study, the selection of the threshold of a detector (γ_k) is done by fixing $\mathcal{P}_F^k$ and $\mathcal{T}_k$. For a batch local detector, this is done via inversion of (6). For the sequential case, this can be achieved via common root-finding methods applied to (27) and (33).

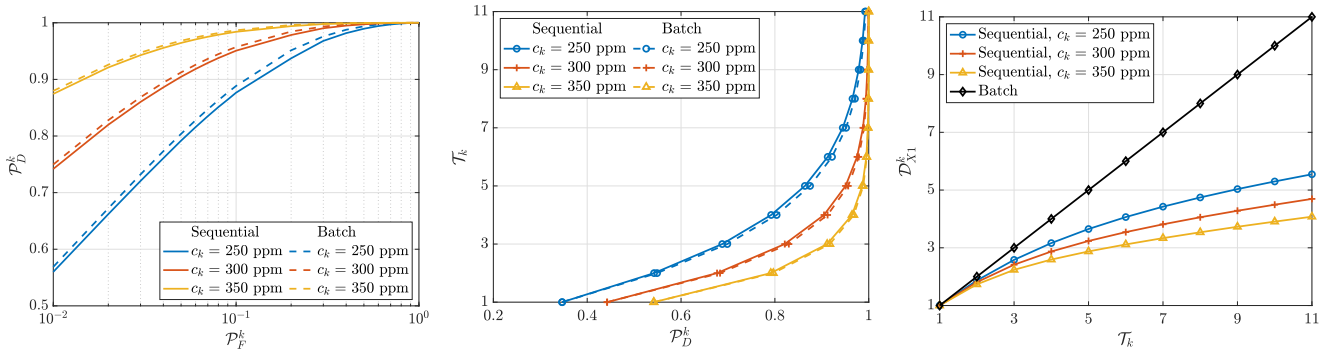
Fig. 3 shows the ROC *surfaces* of the kth sensor in the case of a batch detector (Fig. 3(a)) and a sequential detector (Fig. 3(b) and (c)). These plots are obtained using the relations introduced in Section V. It is immediate to notice how the probability of detection strongly depends on c_k, for a fixed probability of false alarm. In terms of *area under the curve* (AUC), as $c_k \to 0$, we have AUC $\to 0.5$ (random detector), while as $c_k \to \infty$, AUC $\to 1$ (perfect detector), regardless of the used approach.[5] Furthermore, the figure highlights the decision

[5]We remind that, for a generic detector, AUC $\triangleq \int_0^1 \mathcal{P}_D(\mathcal{P}_F)\,\mathrm{d}\mathcal{P}_F$, where $\mathcal{P}_F$ and $\mathcal{P}_D$ are the probability of false alarm and detection, respectively.



(a) Batch detector (colorbar shows $\mathcal{T}_k$) (b) Sequential detector (colorbar shows $\mathcal{D}_{11}^k$) (c) Sequential detector (colorbar shows $\mathcal{D}_{X1}^k$)

Fig. 3. ROC surfaces of local detectors using batch and sequential approach (red line indicates performances at $\mathcal{P}_F^k = 0.05$).



(a) ROC curves of the sensors (b) $\mathcal{P}_D^k$ vs. $\mathcal{T}_k$, with $\mathcal{P}_F^k = 0.05$ (c) $\mathcal{T}_k$ vs. $\mathcal{D}_{X1}^k$, with $\mathcal{P}_F^k = 0.05$

Fig. 4. Detection performances of the sensor.

delays in the two different approaches as the probability of false alarm and c_k change. Fig. 3(a) shows a constant decision delay equal to $\mathcal{T}_k$, while the remaining surfaces highlight the changes in $\mathcal{D}_{11}^k$ and $\mathcal{D}_{X1}^k$. In particular, when $\mathcal{P}_F^k \to 1$, we have $(\mathcal{D}_{11}^k, \mathcal{D}_{X1}^k) \to (1, 1)$, while when $\mathcal{P}_F^k \to 0$, we obtain $(\mathcal{D}_{11}^k, \mathcal{D}_{X1}^k) \to (\mathcal{T}_k, \mathcal{T}_k)$ thanks to the truncation that prevents the delays to diverge to infinity. Finally, in the sequential case, the plots show how the delays tend to lower from $\mathcal{T}_k$ to 1 at a faster rate with respect to $\mathcal{P}_F^k$ as c_k increases.

The comparison between a batch and a sequential detector can be facilitated using Fig. 4. In particular, Fig. 4(a) displays three sets of ROC curves at different values of c_k showing the negligible difference in performance between the two approaches. Meanwhile, Fig. 4(b) shows that once $\mathcal{P}_F^k$ has been fixed, the value of $\mathcal{T}_k$ required to achieve a desired value of $\mathcal{P}_D^k$ is similar in the case of batch and the sequential approach. Hence we can say that the differences in terms of detection accuracy between the batch approach and the sequential approach are negligible. The main advantage of a sequential approach can be seen in Fig. 4(c) where the decision delay $\mathcal{D}_{X1}^k$ is always smaller than $\mathcal{T}_k$ with this difference increasing as we allow higher values of $\mathcal{T}_k$. This highlights that, once the probability of false alarm has been fixed, a sensor can perform detection with a smaller decision delay when a sequential approach is used rather than a batch approach at virtually the same probability of detection.

Fig. 5 shows the values of the ATP using the different architectures once the probability of occurrence of the dispersion $\mathcal{P}_1 \triangleq \mathbb{P}(\mathcal{H}_1)$ is marginalized, making it easier to compare

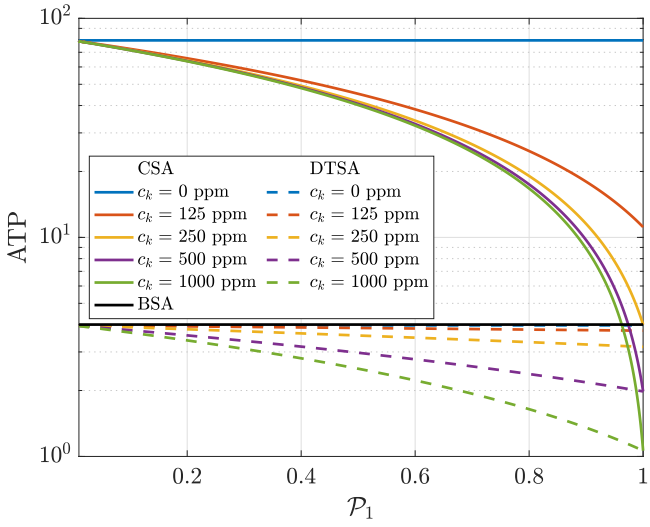
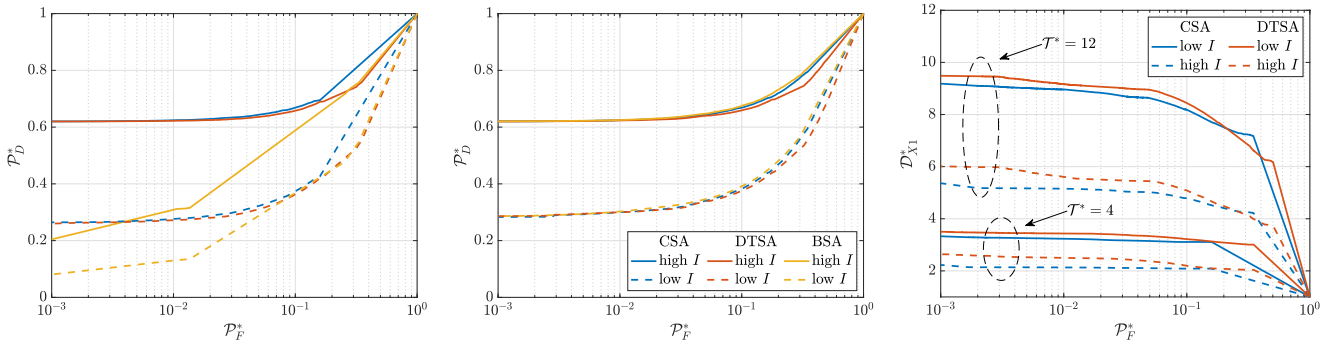


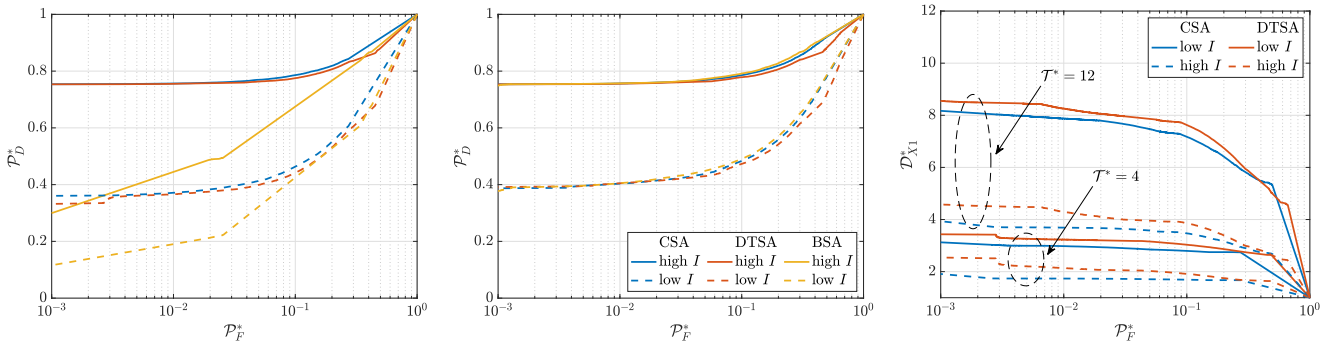
Fig. 5. Average transmission period vs. probability of occurrence.

ATP_{CSA} and ATP_{BSA}. This is because it is fair to assume that such an event happens with low frequency, with the desirable reduction communication in the WSN when $\mathcal{H}_1$ does not occur. We can notice that ATP_{CSA} and ATP_{DTSA} increase as $\mathcal{P}_1$ decreases. However, while ATP_{DTSA} is upper-bounded by ATP_{BSA} (as discussed in Section VI), the behavior of ATP_{CSA} relative to ATP_{BSA} varies according to both $\mathcal{P}_1$ and c_k. In the limit case of $\mathcal{P}_1 = 1$ (resp. $\mathcal{P}_1 = 0$), we can see the values of the



(a) ROC curves, $\mathcal{T}^* = 4$, $K = 9$ (b) ROC curves, $\mathcal{T}^* = 12$, $K = 9$ (c) Decision delay curves, $K = 9$

Fig. 6. ROC curves and decision delay curves, $K = 9$.



(a) ROC curves, $\mathcal{T}^* = 4$, $K = 16$ (b) ROC curves, $\mathcal{T}^* = 12$, $K = 16$ (c) Decision delay curves, $K = 16$

Fig. 7. ROC curves and decision delay curves, $K = 16$.

ATP's in the hypothesis $\mathcal{H}_1$ (resp. $\mathcal{H}_0$): in $\mathcal{H}_1$, ATP_{CSA} tends to increase for lower values of c_k eventually leading to values greater than ATP_{BSA}; in $\mathcal{H}_0$, ATP_{CSA} is sensibly higher than ATP_{BSA}, regardless of c_k. Thus, ATP_{CSA} shows an improvement in the reduction of communication costs when the assumption of low $\mathcal{P}_1$ holds.

Next, we discuss the performances of the FC for each of the four configurations mentioned above in terms of *global probability of false alarm* $\mathcal{P}_F^* \triangleq \mathbb{P}_0(\widehat{\mathcal{H}} = \mathcal{H}_1)$, *global probability of detection* $\mathcal{P}_D^* \triangleq \mathbb{P}_1(\widehat{\mathcal{H}} = \mathcal{H}_1)$, and *global decision delay* (in $\mathcal{H}_1$) defined as $\mathcal{D}_{X1}^* \triangleq \mathbb{E}_1(t^*)$ for the CSA and DTSA, and equal to $\mathcal{D}_{X1}^* \triangleq \mathcal{T}^*$ for the BSA. The results are reported at increasing values of $\mathcal{T}^*$ for comparison purposes.

Figs. 6 and 7 illustrate the ROC curves and the curves where $\mathcal{D}_{X1}$ is shown as function of $\mathcal{P}_F^*$. Different points of the curve are obtained by applying different values of global threshold γ^* to the FC's detection rule. The plots report results for $\mathcal{T}^* \in \{4, 12\}$ (as these are multiples of $\mathcal{T}_k$).[6] The corresponding values of AUC are reported in Table III (with the intermediate scenario with $\mathcal{T}^* = 8$ also present). Analogously to the AUC of the ROC curve, we define $\text{AUC}(\mathcal{D}_{X1}^*) \triangleq \int_0^1 \mathcal{D}_{X1}^*(\mathcal{P}_F^*)\, \mathrm{d}\mathcal{P}_F^*$ to facilitate the discussion of Fig. 7(c) and (f). This metric is the mean value of $\mathcal{D}_{X1}^*$ over the domain of $\mathcal{P}_F^*$ and its values are reported in Table IV.

[6]Higher values of $\mathcal{T}^*$ are not reported as they did not show any significant changes in the ROC curves and in the respective values of AUC.

TABLE III
AUC IN THE SIMULATED CONFIGURATIONS

K	Method	$\mathcal{T}^* = 4$ low I	$\mathcal{T}^* = 4$ high I	$\mathcal{T}^* = 8$ low I	$\mathcal{T}^* = 8$ high I	$\mathcal{T}^* = 12$ low I	$\mathcal{T}^* = 12$ high I
9	CSA	**0.6556**	**0.8173**	0.6672	**0.8252**	0.6754	0.8313
	DTSA	0.6509	0.8121	0.6593	0.8169	0.6654	0.8208
	BSA	0.6307	0.7710	**0.6686**	0.8251	**0.6781**	**0.8343**
16	CSA	**0.7183**	**0.8878**	**0.7343**	**0.8945**	0.7428	0.8988
	DTSA	0.7047	0.8786	0.7157	0.8822	0.7221	0.8848
	BSA	0.6724	0.8294	0.7320	0.8911	**0.7444**	**0.9010**

TABLE IV
$\text{AUC}(\mathcal{D}_{X1}^*)$ IN THE SIMULATED CONFIGURATIONS

K	Method	$\mathcal{T}^* = 4$ low I	$\mathcal{T}^* = 4$ high I	$\mathcal{T}^* = 8$ low I	$\mathcal{T}^* = 8$ high I	$\mathcal{T}^* = 12$ low I	$\mathcal{T}^* = 12$ high I
9	CSA	**2.2263**	**1.6320**	**3.8154**	**2.4724**	**5.3837**	**3.2943**
	DTSA	2.4062	1.7448	3.9661	2.5883	5.5021	3.4148
	BSA	4	4	8	8	12	12
16	CSA	**2.1243**	**1.4326**	**3.4569**	**1.9629**	**4.7146**	**2.4609**
	DTSA	2.3249	1.5427	3.6606	2.0990	4.9415	2.6321
	BSA	4	4	8	8	12	12

The ROC curves show that increasing the number of sensors improves $\mathcal{P}_D^*$, irrespective of the algorithm used. There are two reasons for this: Firstly, a larger number of sensors provides more information to the FC, enabling better discrimination between hypotheses. Secondly, since gas dispersions are anisotropic, having more sensors increases the chances of more sensors being

in contact with the gas plume, resulting in a greater number of sensors experiencing $c_k > 0$, which enables non-random local detections. Another noticeable behavior is the higher value of $\mathcal{P}_D^*$ when the intensity I increases. This is because increasing I (fixing the other parameters) results in a higher c_k for those sensors already in the gas plume, as well as more sensors experiencing $c_k > 0$ (see Fig. 2 for a visual description of the effect of an increase of I). Such behavior of $\mathcal{P}_D^*$ with respect to K and I are numerically confirmed by an increase of AUC.

Using Table III, one can notice an increase in the AUC as higher values of $\mathcal{T}^*$ are used. Moreover, at $\mathcal{T}^* = 4$, the reported values show $\text{AUC}_{\text{CSA}} > \text{AUC}_{\text{DTSA}} > \text{AUC}_{\text{BSA}}$, with the difference in AUC (averaged among the four configurations) between the CSA and BSA, being 0.0439. This changes at $\mathcal{T}^* = 12$, showing a convergence trend in the AUC, with the BSA having the highest values. Nevertheless, the average difference in AUC between the CSA and BSA is equal to -0.0024, making this difference negligible.

Further analysis of the results showed the reason behind the negligible differences in performance obtained by all the architectures as higher values of $\mathcal{T}^*$ are used. This lies in the anisotropic behavior of gas dispersions. In the simulated scenarios, a non-negligible number of Monte Carlo runs resulted in none of the sensors experiencing $c_k > 0$. In such a scenario all the sensors (and so the FC) acts as a random detector regardless of the current value of $\mathcal{T}^*$.

Unlike the previous discussion, the benefits in terms of $\mathcal{D}_{X1}^*$, as K and I increase, are only experienced by the CSA and the DTSA and are shown in Fig. 7(c) and (f). This means that while, on one hand, we reach converging values of AUC by increasing $\mathcal{T}^*$, on the other hand, we are further amplifying the difference in $\text{AUC}(\mathcal{D}_{X1}^*)$ in favor of the sequential algorithms, and in particular the CSA. Table IV clearly shows that $\text{AUC}(\mathcal{D}_{X1}^*)_{\text{CSA}} > \text{AUC}(\mathcal{D}_{X1}^*)_{\text{DTSA}} > \text{AUC}(\mathcal{D}_{X1}^*)_{\text{BSA}}$, for all configurations and values of $\mathcal{T}^*$. This is because, in the CSA and DTSA, $\mathcal{D}_{X1}^*$ grows slower than $\mathcal{T}^*$, unlike in the BSA where the growth is identical.

Both Figs. 6 and 7 show how the selection of the threshold γ^* affects performances. It can be seen how lowering γ^* simultaneously results in a higher $\mathcal{P}_D^*$ and lower $\mathcal{D}_{X1}^*$, with the drawback of an increased value of $\mathcal{P}_F^*$. Nevertheless, the curves show how both the CSA and DTSA are able to have lower $\mathcal{P}_F^*$ maintaining a steady level of $\mathcal{P}_D^*$ and $\mathcal{D}_{X1}^*$. This is especially visible at low values of $\mathcal{T}^*$.

The appropriate value of γ^* can be found via simulation after selecting a metric to satisfy. Possible strategies include: a) given a fixed number of sensors, the threshold is chosen by satisfying a desired maximum $\mathcal{P}_F^*$; b) given a fixed number of sensors, the threshold is chosen so that a minimum value of $\mathcal{P}_D^*$ is achieved given a value of I; c) the threshold is selected by minimizing the *Bayes Risk*; d) the threshold is chosen, together with the number of sensors, so that both $\mathcal{P}_F^*$ and $\mathcal{P}_D^*$ satisfy the desired requirements, given a value of I.

To conclude, the two proposed algorithms present the following differences in terms of performance and complexity:

- The CSA shows superior performances with respect to the DTSA both in terms of detection accuracy and decision delay;
- The CSA shows a great advantage in terms of communication costs, while the DTSA requires more frequent transmissions from the sensors to the FC;
- The DTSA requires less computations since the FC needs to update the detection statistic only when a decision is taken by the sensors. The CSA, on the other hand, requires the FC to update the detection statistic at each instant.

Thus, because of its high performance, the CSA is particularly suitable for highly safety-critical applications like hazardous gas detection. The DTSA, still maintaining high performances, shows a lower degree of accuracy and higher delay with respect to the CSA as well as higher communication costs. However, the DTSA's lower requirement in terms of computations performed by the FC makes it a desirable solution as long as a higher number of sensors is employed.

VIII. Conclusions and Future Directions

We proposed two sequential algorithms addressing the task of distributed gas detection in WSNs, named CSA and DTSA. The setup consists of sensors taking binary decisions via SD and transmitting them to a FC which takes a final decision benefiting from the integration of meteorological data and the dispersion model. The proposed methods constitute fully sequential alternatives to the traditional batch approach (BSA), with the further innovation introduced by the CSA of a time-aware sequential fusion. This enabled a significant improvement in terms of detection accuracy and delay, especially desired in such a time-critical application. System performance was also assessed in terms of communication costs showing how a time-aware algorithm as the CSA greatly reduces transmissions from sensors to the FC. The case study of CO_2 dispersion confirmed the validity of the proposed architectures.

Future works include a) the reduction of complexity via more efficient strategies for the searching of $(\boldsymbol{\theta}, I)$, including the estimation of possible variations of I over time; b) modeling erroneous communication channels; c) use of Bayesian methods to improve detection and parameter estimation; d) more accurate statistical characterization of the signal measured by the sensors including possible correlations between measurements in space and time; e) development of algorithms accounting for imperfect knowledge of the dispersion model; f) use of more comprehensive dispersion models, or direct use of computational fluid dynamics software; g) integration of machine learning strategies for improved detection performances.

Appendix A
Covariance Matrices

The following is the derivation of the matrices $\boldsymbol{\Sigma}_{\boldsymbol{z}_i^k}$ and $\boldsymbol{\Sigma}_{\boldsymbol{v}_i^k}$, for any $k = 1, \ldots, K$ and $i = 1, \ldots, \mathcal{T}_k$. The element of the matrix $\boldsymbol{\Sigma}_{\boldsymbol{v}_i^k}$ located in the rth row and sth column is defined as the following:

$$\left[\boldsymbol{\Sigma}_{\boldsymbol{v}_i^k}\right]_{r,s} \triangleq \mathbb{C}\text{ov}\left(\Lambda_r^k - \gamma_k, \Lambda_s^k - \gamma_k\right) = \mathbb{C}\text{ov}\left(\Lambda_r^k, \Lambda_s^k\right) .$$

When $r = s \leq i$, we have that:

$$\left[\boldsymbol{\Sigma}_{\boldsymbol{v}_i^k}\right]_{r,s} = \mathbb{V}\text{ar}\left(\Lambda_s^k\right) = s\sigma_k^2 = r\sigma_k^2 .$$

On the other hand, when $r < s \leq i$:

$$\begin{aligned}\left[\boldsymbol{\Sigma}_{\boldsymbol{v}_i^k}\right]_{r,s} &= \mathbb{C}\mathrm{ov}\left(\Lambda_r^k, \Lambda_r^k + \sum_{j=r+1}^{s} \lambda_j^k\right) \\ &= \mathbb{E}\left(\Lambda_r^k\left(\Lambda_r^k + \sum_{j=r+1}^{s} \lambda_j^k\right)\right) \\ &\quad - \mathbb{E}\big(\Lambda_r^k\big)\mathbb{E}\left(\Lambda_r^k + \sum_{j=r+1}^{s} \lambda_j^k\right) \\ &= \mathbb{E}\Big(\big(\Lambda_r^k\big)^2\Big) - \mathbb{E}^2\big(\Lambda_r^k\big) = \mathbb{V}\mathrm{ar}\big(\Lambda_r^k\big) = r\sigma_k^2 \,.\end{aligned}$$

Analogously, when $s < r \leq i$, $\left[\boldsymbol{\Sigma}_{\boldsymbol{v}_i^k}\right]_{r,s} = s\sigma_k^2$. Hence, it is easy to obtain the following:

$$\left[\boldsymbol{\Sigma}_{\boldsymbol{v}_i^k}\right]_{r,s} = \min\{r,s\}\cdot\sigma_k^2 \,, \quad \forall\, r \leq i, s \leq i \,.$$

For the case of $\boldsymbol{\Sigma}_{\boldsymbol{z}_i^k}$, the previous holds as long as $r < i$ and $s < i$. In fact, when $r < s = i$:

$$\begin{aligned}\left[\boldsymbol{\Sigma}_{\boldsymbol{v}_i^k}\right]_{r,s} &\triangleq \mathbb{C}\mathrm{ov}\big(\Lambda_r^k - \gamma_k, -\Lambda_s^k + \gamma_k\big) = \mathbb{C}\mathrm{ov}\big(\Lambda_r^k, -\Lambda_i^k\big) \\ &= \mathbb{C}\mathrm{ov}\left(\Lambda_r^k, -\Lambda_r^k - \sum_{j=r+1}^{i} \lambda_j^k\right) \\ &= \mathbb{E}\left(-\Lambda_r^k\left(\Lambda_r^k + \sum_{j=r+1}^{i} \lambda_j^k\right)\right) \\ &\quad - \mathbb{E}\big(\Lambda_r^k\big)\mathbb{E}\left(-\Lambda_r^k - \sum_{j=r+1}^{i} \lambda_j^k\right) \\ &= -\mathbb{E}\Big(\big(\Lambda_r^k\big)^2\Big) + \mathbb{E}^2\big(\Lambda_r^k\big) = -\mathbb{V}\mathrm{ar}\big(\Lambda_r^k\big) = -r\sigma_k^2 \,.\end{aligned}$$

Similarly, when $s < r = i$, $\left[\boldsymbol{\Sigma}_{\boldsymbol{v}_i^k}\right]_{r,s} = -s\sigma_k^2$. Lastly, when $r = s = i$, we have:

$$\begin{aligned}\left[\boldsymbol{\Sigma}_{\boldsymbol{z}_i^k}\right]_{r,s} &\triangleq \mathbb{C}\mathrm{ov}\big(-\Lambda_r^k + \gamma_k, -\Lambda_s^k + \gamma_k\big) = \mathbb{V}\mathrm{ar}\big(-\Lambda_i^k\big) \\ &= \mathbb{V}\mathrm{ar}\big(\Lambda_i^k\big) = i\sigma_k^2 \,.\end{aligned}$$

These four cases form the following rule:

$$\begin{aligned}\left[\boldsymbol{\Sigma}_{\boldsymbol{z}_i^k}\right]_{r,s} &= \begin{cases} \min\{r,s\}\cdot\sigma_k^2 \,, & \text{if } r < i \text{ and } s < i \\ -r\sigma_k^2 \,, & \text{if } r < s = i \\ -s\sigma_k^2 \,, & \text{if } s < r = i \\ i\sigma_k^2 \,, & \text{if } r = s = i \end{cases} \\ &= \sigma_k^2 \min\{r,s\}\big[1 + 2\delta_{i,\max\{r,s\}}\big(\delta_{i,\min\{r,s\}} - 1\big)\big] \,.\end{aligned}$$

Appendix B
Average Transmission Period in the CSA

We here report the derivation of the $\mathrm{ATP}_{\mathrm{CSA}}$ reported in (37). We show the proof for the case where $\mathcal{H}_0$ is true:

$$\begin{aligned}&\mathbb{E}_0\left(t_b^k - t_a^k \middle| \sum_{t=t_a^k}^{t_b^k} \tau_t^k = 2\right) = \mathbb{E}_0\left(t_i^k \middle| \sum_{t=1}^{t_i^k} \tau_t^k = 1\right) \\ &= \sum_{i=1}^{\infty} \mathbb{E}_0\left(t_i^k \middle| \sum_{t=1}^{t_i^k} \tau_t^k = 1\right) \mathbb{P}_0\left(\sum_{t=1}^{t_i^k} \tau_t^k = 1\right) \\ &= \sum_{i=1}^{\infty} \big[(i-1)\mathbb{E}_0\big(t_1^k \big| d_1^k = \mathcal{H}_0\big) + \mathbb{E}_0\big(t_1^k \big| d_i^k = \mathcal{H}_1\big)\big] \\ &\quad \times \mathbb{P}_0\big(d_1^k = \mathcal{H}_1\big)\big[1 - \mathbb{P}_0\big(d_i^k = \mathcal{H}_1\big)\big]^{i-1} \\ &= \sum_{i=1}^{\infty} \big[(i-1)\mathcal{T}_k + \mathcal{D}_{10}^k\big]\mathcal{P}_F^k\big(1 - \mathcal{P}_F^k\big)^{i-1} \\ &= \mathcal{P}_F^k \sum_{i=0}^{\infty} \big(i\,\mathcal{T}_k + \mathcal{D}_{10}^k\big)\big(1 - \mathcal{P}_F^k\big)^{i} \\ &= \mathcal{P}_F^k\left(\mathcal{T}_k \sum_{i=0}^{\infty} i\big(1 - \mathcal{P}_F^k\big)^{i} + \mathcal{D}_{10}^k \sum_{i=0}^{\infty} \big(1 - \mathcal{P}_F^k\big)^{i}\right) \\ &= \mathcal{P}_F^k\left(\mathcal{T}_k \frac{1 - \mathcal{P}_F^k}{\big(\mathcal{P}_F^k\big)^2} + \frac{\mathcal{D}_{10}^k}{\mathcal{P}_F^k}\right) = \frac{\mathcal{T}_k(1 - \mathcal{P}_F^k) + \mathcal{P}_F^k\mathcal{D}_{10}^k}{\mathcal{P}_F^k} \\ &= \frac{\mathcal{D}_{X0}^k}{\mathcal{P}_F^k}.\end{aligned}$$

Similarly, we obtain $\mathcal{D}_{X1}^k(c_k)/\mathcal{P}_D^k(c_k)$ when $\mathcal{H}_1$ is true.

Acknowledgment

This research is a part of BRU21 NTNU Research and Innovation Program on Digital and Automation Solutions for the Oil and Gas Industry (www.ntnu.edu/bru21).

References

[1] G. Tabella, Y. D. Martino, D. Ciuonzo, N. Paltrinieri, X. Wang, and P. Salvo Rossi, "Decision fusion for carbon dioxide release detection from pressure relief devices," in *Proc. IEEE 12th Sens. Array Multichannel Signal Process. Workshop*, 2022, pp. 46–50.

[2] G. Tabella, Y. Di Martino, D. Ciuonzo, N. Paltrinieri, X. Wang, and P. Salvo Rossi, "Sensor fusion for detection and localization of carbon dioxide releases for Industry 4.0," in *Proc. IEEE 25th Int. Conf. Inf. Fusion*, 2022, pp. 1–8.

[3] S. Al-Sarawi, M. Anbar, R. Abdullah, and A. B. Al Hawari, "Internet of Things market analysis forecasts, 2020–2030," in *Proc. IEEE 4th World Conf. Smart Trends Syst., Secur. Sustainability*, 2020, pp. 449–453.

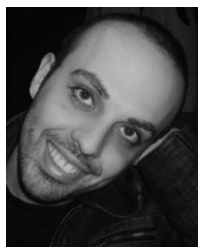
[4] R. Abielmona, R. Falcon, N. Zincir-Heywood, and H. A. Abbass, Eds., *Recent Advance in Computational Intelligence in Defense and Security* (Studies in Computational Intelligence Series). Cham, Switzerland: Springer, 2016.

[5] G. Tabella, N. Paltrinieri, V. Cozzani, and P. Salvo Rossi, "Wireless sensor networks for detection and localization of subsea oil leakages," *IEEE Sens. J.*, vol. 21, no. 9, pp. 10890–10904, May 2021.

[6] G. Tabella, D. Ciuonzo, N. Paltrinieri, and P. Salvo Rossi, "Spatio-temporal decision fusion for quickest fault detection within industrial plants: The oil and gas scenario," in *Proc. IEEE 24th Int. Conf. Inf. Fusion*, 2021, pp. 1–8.

[7] B. Ristic, A. Gunatilaka, and R. Gailis, "Localisation of a source of hazardous substance dispersion using binary measurements," *Atmosph. Environ.*, vol. 142, pp. 114–119, 2016.
[8] D. D. Selvaratnam, I. Shames, D. V. Dimarogonas, J. H. Manton, and B. Ristic, "Co-operative estimation for source localisation using binary sensors," in *Proc. IEEE 56th Annu. Conf. Decis. Control*, 2017, pp. 1572–1577.
[9] T. Sahoo, *Process Plants-Shutdown and Turnaround Management*. Boca Raton, FL, USA: Taylor & Francis Group, 2014.
[10] T. Zhao and A. Nehorai, "Detecting and estimating biochemical dispersion of a moving source in a semi-infinite medium," *IEEE Trans. Signal Process.*, vol. 54, no. 6, pp. 2213–2225, Jun. 2006.
[11] S. Aldalahmeh, M. Ghogho, and A. Swami, "Fast distributed detection, localization, and estimation of a diffusive target in wireless sensor networks," in *Proc. IEEE 7th Int. Symp. Wireless Commun. Syst.*, 2010, pp. 882–886.
[12] B. Ristic, A. Gunatilaka, and R. Gailis, "Achievable accuracy in gaussian plume parameter estimation using a network of binary sensors," *Inf. Fusion*, vol. 25, pp. 42–48, 2015.
[13] S. Vijayakumaran, Y. Levinbook, and T. F. Wong, "Maximum likelihood localization of a diffusive point source using binary observations," *IEEE Trans. Signal Process.*, vol. 55, no. 2, pp. 665–676, Feb. 2007.
[14] A. Gunatilaka, B. Ristic, A. Skvortsov, and M. Morelande, "Parameter estimation of a continuous chemical plume source," in *Proc. IEEE 11th Int. Conf. Inf. Fusion*, 2008, pp. 1–8.
[15] J. Shi et al., "Real-time plume tracking using transfer learning approach," *Comput. Chem. Eng*, vol. 172, 2023, Art. no. 108172.
[16] R. R. Tenney and N. R. Sandell, "Detection with distributed sensors," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-17, no. 4, pp. 501–510, Jul. 1981.
[17] I. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *IEEE Commun. Mag.*, vol. 40, no. 8, pp. 102–114, Aug. 2002.
[18] J.-F. Chamberland and V. Veeravalli, "Decentralized detection in sensor networks," *IEEE Trans. Signal Process.*, vol. 51, no. 2, pp. 407–416, Feb. 2003.
[19] B. Chen, L. Tong, and P. K. Varshney, "Channel-aware distributed detection in wireless sensor networks," *IEEE Signal Process. Mag.*, vol. 23, no. 4, pp. 16–26, Jul. 2006.
[20] D. Ciuonzo and P. Salvo Rossi, "Distributed detection of a non-cooperative target via generalized locally-optimum approaches," *Inf. Fusion*, vol. 36, pp. 261–274, 2017.
[21] A. Wald, *Sequential Analysis* (Wiley publication in mathematical statistics Series). New York, NY, USA: Wiley, 1947.
[22] A. Wald and J. Wolfowitz, "Optimum character of the sequential probability ratio test," *Ann. Math. Statist.*, vol. 19, no. 3, pp. 326–339, 1948.
[23] H. Poor, *An Introduction to Signal Detection and Estimation* (Texts in Electrical Engineering Series), 2nd ed. New York, NY, USA: Springer, 2013.
[24] D. Siegmund, *Sequential Analysis: Tests and Confidence Intervals* (Springer Series in Statistics). New York, NY, USA: Springer, 1985.
[25] Y. Mei, "Asymptotic optimality theory for decentralized sequential hypothesis testing in sensor networks," *IEEE Trans. Inf. Theory*, vol. 54, no. 5, pp. 2072–2089, May 2008.
[26] G. Fellouris and G. V. Moustakides, "Decentralized sequential hypothesis testing using asynchronous communication," *IEEE Trans. Inf. Theory*, vol. 57, no. 1, pp. 534–548, Jan. 2011.
[27] Y. Yilmaz, G. V. Moustakides, and X. Wang, "Cooperative sequential spectrum sensing based on level-triggered sampling," *IEEE Trans. Signal Process.*, vol. 60, no. 9, pp. 4509–4524, Sep. 2012.
[28] Y. Yilmaz, G. V. Moustakides, and X. Wang, "Channel-aware decentralized detection via level-triggered sampling," *IEEE Trans. Signal Process.*, vol. 61, no. 2, pp. 300–315, Jan. 2013.
[29] Y. Yilmaz and X. Wang, "Sequential distributed detection in energy-constrained wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 62, no. 12, pp. 3180–3193, Jun. 2014.
[30] S. Li, X. Li, X. Wang, and J. Liu, "Decentralized sequential composite hypothesis test based on one-bit communication," *IEEE Trans. Inf. Theory*, vol. 63, no. 6, pp. 3405–3424, Jun. 2017.
[31] S. Tantaratana and H. Poor, "Asymptotic efficiencies of truncated sequential tests," *IEEE Trans. Inf. Theory*, vol. 28, no. 6, pp. 911–923, Nov. 1982.
[32] M. Guerriero, V. Pozdnyakov, J. Glaz, and P. Willett, "A repeated significance test with applications to sequential detection in sensor networks," *IEEE Trans. Signal Process.*, vol. 58, no. 7, pp. 3426–3435, Jul. 2010.
[33] P. Khanduri, D. Pastor, V. Sharma, and P. K. Varshney, "Truncated sequential non-parametric hypothesis testing based on random distortion testing," *IEEE Trans. Signal Process.*, vol. 67, no. 15, pp. 4027–4042, Aug. 2019.
[34] L. Hu, J. Zhang, X. Wang, S. Wang, and E. Zhang, "Decentralized truncated one-sided sequential detection of a noncooperative moving target," *IEEE Signal Process. Lett.*, vol. 25, no. 10, pp. 1490–1494, Oct. 2018.
[35] X. Cheng, D. Ciuonzo, P. Salvo Rossi, X. Wang, and W. Wang, "Multi-bit & sequential decentralized detection of a noncooperative moving target through a generalized Rao test," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 7, pp. 740–753, 2021.
[36] H. Darvishi, D. Ciuonzo, and P. Salvo Rossi, "A machine-learning architecture for sensor fault detection, isolation, and accommodation in digital twins," *IEEE Sensors J.*, vol. 23, no. 3, pp. 2522–2538, Feb. 2023.
[37] Y. Liu, Z. Pang, M. Karlsson, and S. Gong, "Anomaly detection based on machine learning in IoT-based vertical plant wall for indoor climate control," *Building Environ.*, vol. 183, 2020, Art. no. 107212.
[38] A. D'Costa, V. Ramachandran, and A. M. Sayeed, "Distributed classification of gaussian space-time sources in wireless sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 22, no. 6, pp. 1026–1036, Aug. 2004.
[39] R. E. Britter and J. McQuaid, *Workbook on the Dispersion of Dense Gases*. Bootle, U.K.: Health and Safety Executive, 1988.
[40] TNO, *Yellow Book — Methods for the Calculation of Physical Effects*. The Hague, The Netherlands: The Committee Prevention Disasters Hazardous Mater., 2005.
[41] D. A. Crowl and J. F. Louvar, *Chemical Process Safety: Fundamentals With Applications*, 4th ed. London, U.K.: Pearson Educ., 2019.
[42] S. Mannan, *Lees' Loss Prevention in the Process Industries*, 4th ed. Oxford, U.K.: Butterworth-Heinemann, 2012.
[43] NIST, "Thermophysical properties of carbon dioxide," 2023. [Online]. Available: https://webbook.nist.gov/cgi/fluid.cgi?ID=C124389&Action=Page
[44] J. M. Smith, H. C. Van Ness, M. M. Abbott, and M. T. Swihart, *Introduction to Chemical Engineering Thermodynamics*, 8th ed. New York, NY, USA: McGraw-Hill Educ., 2018.

Gianluca Tabella (Graduate Student Member, IEEE) was born in Suzzara, Italy, in 1993. He received the B.Sc. degree in chemical and biochemical engineering (with specialization in process engineering) and the M.Sc. degree in chemical and process engineering (with specialization in offshore engineering) from the University of Bologna, Bologna, Italy, in 2017 and 2019, respectively. Since 2020, he has been working toward the Ph.D. degree in electronics and telecommunication with the Department of Electronic Systems, NTNU Norwegian University of Science and Technology, Trondheim, Norway. He was a Visiting Scholar with Columbia University, New York, NY, USA, in 2022. His research interests include distributed detection and localization with a focus on industrial and oil&gas applications.



Domenico Ciuonzo (Senior Member, IEEE) received the Ph.D. degree from the University of Campania "L. Vanvitelli", Naples, Italy. He is currently a Tenure-track Assistant Professor with the University of Naples "Federico II", Naples. Since 2011, he has been holding several visiting researcher appointments, such as NATO CMRE, UConn, NTNU, and CTTC. His research interests include data fusion, statistical signal processing, wireless sensor networks, IoT, and machine learning. He was the recipient of the two Best Paper Awards, such as IEEE ICCCS 2019 and Elsevier Computer Networks 2020, 2019 Exceptional Service award from IEEE AESS, 2020 Early-Career Technical Achievement Award from IEEE SENSORS COUNCIL for sensor networks/systems, and 2021 Early-Career Award from IEEE AESS for contributions to decentralized inference and sensor fusion in networked sensor systems.

Yasin Yilmaz (Senior Member, IEEE) received the Ph.D. degree in electrical engineering from Columbia University, New York, NY, USA, in 2014. He is currently an Associate Professor of electrical engineering with the University of South Florida, Tampa, FL, USA. His research interests include machine learning, statistical signal processing, and their applications to computer vision, cybersecurity, biomedical systems, energy systems, transportation systems, and communication systems.

Xiaodong Wang (Fellow, IEEE) received the Ph.D. degree in electrical engineering from Princeton University, Princeton, NJ, USA. He is currently a Professor of electrical engineering with Columbia University, New York, NY, USA. Among his publications is a book entitled *Wireless Communication Systems: Advanced Techniques for Signal Reception*, published by Prentice Hall in 2003. He has authored or coauthored in his research areas, which include general areas of computing, signal processing, communications, wireless communications, statistical signal processing, and genomic signal processing.

Dr. Wang was the recipient of the 1999 NSF CAREER Award, 2001 IEEE Communications Society and Information Theory Society Joint Paper Award, and 2011 IEEE Communication Society Award for Outstanding Paper on New Communication Topics. He was an Associate Editor for IEEE TRANSACTIONS ON COMMUNICATIONS, IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, IEEE TRANSACTIONS ON SIGNAL PROCESSING, and IEEE TRANSACTIONS ON INFORMATION THEORY. He is listed as an ISI Highly-cited Author.

Pierluigi Salvo Rossi (Senior Member, IEEE) was born in Naples, Italy, in 1977. He received the Dr.Eng. degree (*summa cum laude*) in telecommunications engineering and the Ph.D. degree in computer engineering from the University of Naples "Federico II", Naples, Italy, in 2002 and 2005, respectively. He is currently a Full Professor and the Deputy Head with the Department of Electronic Systems, Norwegian University of Science and Technology, Trondheim, Norway. He is also a part-time Research Scientist with the Department of Gas Technology, SINTEF Energy Research, Trondheim.

He was with the University of Naples "Federico II" with the Second University of Naples, NTNU, Trondheim, and Kongsberg Digital AS, Horten, Norway. He held visiting appointments with Drexel University, Philadelphia, PA, USA, Lund University, Lund, Sweden, NTNU, and Uppsala University, Uppsala, Sweden.

His research interests include the areas of communication theory, data fusion, machine learning, and signal processing. Dr. Salvo Rossi was awarded as an Exemplary Senior Editor of IEEE COMMUNICATIONS LETTERS in 2018. He is/was with the Editorial Board of IEEE SENSORS JOURNAL, IEEE OPEN JOURNAL OF THE COMMUNICATIONS SOCIETY, IEEE TRANSACTIONS ON SIGNAL AND INFORMATION PROCESSING OVER NETWORKS, IEEE COMMUNICATIONS LETTERS, and IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS.