# Distributed sensor failure detection in sensor networks

Tamara Tošić [a,*], Nikolaos Thomos [a,b], Pascal Frossard [a]

[a] Ecole Polytechnique Fédérale de Lausanne (EPFL), Signal Processing Laboratory (LTS4), Lausanne 1015, Switzerland
[b] University of Bern, Communication and Distributed Systems laboratory (CDS), Bern, Switzerland

## ABSTRACT

We investigate the problem of distributed sensors' failure detection in networks with a small number of defective sensors, whose measurements differ significantly from the neighbor measurements. We build on the sparse nature of the binary sensor failure signals to propose a novel distributed detection algorithm based on gossip mechanisms and on Group Testing (GT), where the latter has been used so far in centralized detection problems. The new distributed GT algorithm estimates the set of scattered defective sensors with a low complexity distance decoder from a small number of linearly independent binary messages exchanged by the sensors. We first consider networks with one defective sensor and determine the minimal number of linearly independent messages needed for its detection with high probability. We then extend our study to the multiple defective sensors detection by modifying appropriately the message exchange protocol and the decoding procedure. We show that, for small and medium sized networks, the number of messages required for successful detection is actually smaller than the minimal number computed theoretically. Finally, simulations demonstrate that the proposed method outperforms methods based on random walks in terms of both detection performance and convergence rate.

© 2012 Elsevier B.V. All rights reserved.

## 1. Introduction

Over the past years we have witnessed the emergence of simple and low cost sensors. This has led to wide deployment of robust sensor networks for monitoring signals in numerous applications, for example in medical applications or natural hazard detection. Most of the previous detection methods are analyzed in centralized settings, where one or more nodes have the additional functionality for directing the failure detection. In practice, sensor networks have a dynamic architecture with loose coordination in order to minimize the communication costs. This raises a demand for novel distributed data processing algorithms that are effective under network topology and communication constraints. In particular, it is important to perform failure detection in a distributed way to decrease high communication costs of the centralized algorithm.

In this paper, we propose a novel distributed sensors' failure detection method that employs a simple distance decoder for sparse and binary signals. We consider a sensor network that is represented as a connected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where vertices $\mathcal{V} = \{s_i\}_{i=1}^{S}$ stand for the $S$ sensors in the network and edges $\mathcal{E}$ determine sensors' connectivity (see Fig. 1). For instance, if two sensors $s_i$ and $s_j$ lie within each other's communication range, the edge $e_{i,j} \in \mathcal{E}$ has a non-zero value. We assume that the function measured by sensors is smooth, so the neighbor sensors typically have similar measurements as long as sensors work correctly. At most $K$ out of $S$ sensors are defective, where $K \ll S$ and the defective sensors are scattered in the network. Therefore, the defective sensor identification problem boils down to a sparse binary signal recovery, where non-zero signal values correspond to defective sensors. Our novel distributed detection approach builds on ideas used in GT methods [1] that are commonly applied for centralized systems. The

* Corresponding author. Tel.: +41 21 693 27 08; fax: +41 21 693 76 00.
*E-mail addresses:* tamara.tosic@epfl.ch (T. Tošić),
nikolaos.thomos@epfl.ch (N. Thomos),
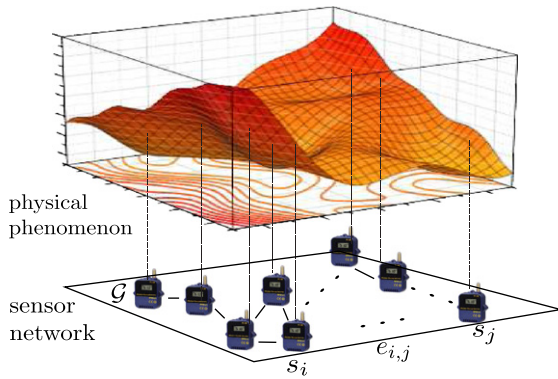pascal.frossard@epfl.ch (P. Frossard).

**Fig. 1.** Ad hoc sensor network measuring a smooth physical phenomenon.

core idea is to perform low-cost experiments in the network, called *tests*, in order to detect the defective sensors. The tests (detection) are performed on pools of sensors by a set of sensors called *master* sensors. Each master sensor compares the sensor measurements based on a similarity measure (e.g., threshold) to detect the presence of defective sensors in its vicinity. The result of this test takes a binary value, which might be possibly altered by noise. Probabilistically designed tests and their outputs at master sensors together build *messages*. They are communicated between the master nodes and their direct neighbors. Next, the messages are disseminated in the network with a gossip algorithm (rumor mongering) [2] that follows a pull protocol [3], where every sensor randomly picks a sender amongst its direct neighbors. Each time a new message reaches a sensor, its value is linearly combined with the current message at this sensor. This increases the diversity of information in the network, which is important for an accurate decoding. The message design and dissemination phases are repeated for several rounds. The specific probabilistic test design described later in this paper ensures the successful detection with high probability with a simple distance decoder (e.g., Hamming decoder), as long as the number of messages exceeds a given threshold. Due to the distributed nature of our algorithm, the detection can be performed in any sensor and not only at a master node. We analyze the detection failure bounds and derive the conditions needed for successful failure detection in the case of a single defective sensor. Then, we provide the error bounds for detection of multiple defective sensors in the network. We show that the number of linearly independent messages required for detection is smaller in practice than given by the theoretical bounds obtained in our worst case analysis. We finally provide simulation results for regular and irregular networks. The experiments outline the advantages of the proposed detection method compared to other binary signal detection algorithms based on the random walk measurements gathering. Our algorithm outperforms random walk detection methods both in terms of the detection accuracy and the convergence rate, because it creates the innovative messages with the higher rate and disseminates them faster.

This paper is organized as follows. Section 2 overviews related works in the detection literature and Section 3 reviews the centralized Group Testing

framework. Section 4 proposes a novel distributed detection method. It describes the message formation and dissemination processes in sensor networks and discusses the detection problem for single and multiple defective sensors. Section 5 presents the simulation results.

## 2. Related works

In general, the detection literature can be mostly classified into fully centralized and semi-distributed methods. In the latter, nodes are grouped into static clusters that independently perform failure detection. Detection algorithms generally employ statistical decoders [4] and often assume that the signal probability distribution and the network connectivity are known. For example, a Bayesian approach in [5] proposes to compute a detection score for a priori defined sets of hypothesis, which depends on the received messages. The hypothesis with the highest score drives the final decision. The binary event detection problem for hierarchically clustered networks is proposed in [6] where the cluster decisions are fused centrally to make a final decision. Surveys on similar methods can be found in [7,8]. For more details on malicious detection, interested readers are referred to the detailed survey [9] and references within.

Another family of detection methods is based on Group Testing (GT) ideas, which originates from the work [1] that targets detection in medical applications. In particular, it proposes a simple idea of pooling blood samples to observe the viral presence in a set, instead of performing tests on every single blood sample separately. Typically, the main target is to minimize the number of tests required to identify all the infected samples, while keeping the detection procedure as simple as possible. GT has been studied more recently in the context of sensor networks for detection of malicious events. In summary, the GT literature can be divided into two main algorithm types, combinatorial and probabilistic. Both approaches aim to minimize the number of tests for detection. The combinatorial GT [10] uses a deterministic test design, while the probabilistic GT applies the knowledge of probability distribution of defective elements for this purpose. From the defect type perspective, detection approaches differ in scenarios with errors, inhibitors or combinations of them and they can be rather naive [11]. In some works detection is performed by iterative elimination of identified non-defective items from the test outcomes. The detection time is typically of order $\mathcal{O}(SB)$, where $B$ is the number of tests and $S$ is the total number of sensors. Particular test design methods improve the effective time for detection in centralized systems. For example, a useful test matrix property called $K$-disjunctness (i.e., the Boolean sum of every $K$ columns does not result in any other column), speeds up the decoding process. This property is used in code designs, e.g., for superimposed codes [12,13] or in detection of malicious users [14]. Finally, a random efficient construction of disjunct matrices is proposed in [15] with a decoding time of $\mathcal{O}(\pi(B) \cdot B \log^2 B + \mathcal{O}(B^2))$, where $B = \mathcal{O}(K^2 \log S)$ is the number of tests and $\pi(\cdot)$ denotes a polynomial. In our knowledge,

this represents the state-of-the-art decoding performance for centralized detection of an arbitrary sparse signal.

Finally, even if the test designs are contingent to the communication limitations in sensor networks, not many works have considered connectivity constraints imposed by the network topology in GT methods. The authors in [16] propose to form tests by a random walk process on well-connected graphs. The minimal number of tests required for detection in this case depends on the random walk mixing time. A bipartite graph structure is considered in [17] with a two-stage hybrid detection method. Here, a subset of defective items in the first stage is determined by pre-designed tests, while the remaining items are tested individually in the next step. Data retrieval for topology-adaptive GT is studied in [18] where a binary tree splitting algorithm is proposed. The above methods however use centralized decision algorithms which are not appropriate for large-scale sensor networks or networks with a dynamic topology because of the high communication costs.

Not many works actually address the detection problem in a distributed way. Moreover, we should emphasize that the term "distributed" in earlier detection works generally refers to the information transmission process between neighboring nodes towards the fusion center that performs detection. In our work, we however consider the problem of distributed detection, where each sensor in the network is able to perform detection task. In other words, motivated by works that address distributed signal processing tasks (e.g., [19]), we target a framework where nodes both transmit and process data in a distributed manner. Our paper specifically focuses on a resource effective distributed failure detection algorithm, which can be deployed in robust monitoring networks. To the best of our knowledge, no former analysis has been proposed for distributed detection methods of sparse binary test signals as proposed in this paper. The proposed algorithm may be used in general distributed sparse detection problems, like for instance, for determining data outliers or abnormalities in networks.

## 3. Centralized detection with probabilistic group testing

We first review the centralized detection of sensor failures with methods based on GT and then describe our novel distributed detection method. We adopt the following notation. Script letters denote sets, $|\cdot|$ represents the number of elements in a set and the $i$-th column and row of the matrix $\mathbf{W}$ are represented with $\mathbf{W}_{:,i}$ and $\mathbf{W}_{i,:}$, respectively.

GT aims at detecting defective items in the set based on the outcome of binary tests. Non-zero entries of an $S$-dimensional binary vector $\mathbf{f} \in \mathbb{F}_2^S$ indicate the defective sensors. $\mathbb{F}_2$ is a finite field of size two and $\mathbf{f}$ is a $K$-sparse signal, where $K \ll S$ and defective sensors are scattered in the network. The tests preformed on sensor measurements are represented with a $B \times S$ dimensional matrix $\mathbf{W}$, where $B$ stands for the number of tests. The non-zero entries of $\mathbf{W}_{i,:} \in \mathbb{F}_2^S$ refer to sensors that participate in the $i$-th test. The Boolean matrix multiplication operator is denoted with $\otimes$. Then, the binary test results are denoted with the test outcome vector $\mathbf{g} = \mathbf{W} \otimes \mathbf{f}$, where $\mathbf{g} \in \mathbb{F}_2^B$.

The design of the matrix $\mathbf{W}$ is crucial for reducing the number of required tests for the detection of defective sensors. This design resembles the design of generator matrices of LDPC codes [20]. Motivated by this similarity, the test matrix $\mathbf{W}$ is constructed as [21]

$$W_{i,j} = \begin{cases} 1 & \text{with probability } q, \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

The sensor participates in a test with a probability $q$. Such a test matrix design assures that it is *disjunct* with high probability [21, Section IV, Definition 1]. In other words, a matrix $\mathbf{W}$ is called $K$-disjunct if no column $\mathbf{W}_{:,i}$ of $\mathbf{W}$ lies in the sub-space formed by any set of $K$ columns $\mathbf{W}_{:,j}$ with $j \neq i$. This property enables detection [21, Section IV, Proposition 2] with a distance decoder (i.e., Hamming distance). The disjunct matrix parameter $\epsilon$ represents the distance decoder threshold for detection. We define the support of the vector with the operator $supp(\cdot)$. For any column $\mathbf{W}_{:,i}$ of the test matrix $\mathbf{W}$ that is $(K,\epsilon)$-disjunct, the decoder verifies if

$$\left| supp(\mathbf{W}_{:,i}) \backslash supp(\mathbf{g}) \right| \leq \epsilon. \tag{2}$$

In other words, the decoder counts the number of positions in the column $\mathbf{W}_{:,i}$ that are different to the outcome vector $\mathbf{g}$. To remind, the non-zero values of the vector $\mathbf{f}$ denote the defective elements. Then, the elements of $\mathbf{f}$ are inferred as non-zero (marked as defective) iff the inequality (2) holds. In [21, Theorem 4], the required number of measurements for successful decoding is equal to $B = \mathcal{O}(K \log(S)/p^3)$ for centralized detection in noisy settings and any single set of defective sensors. The noise alternates the non-zero entries in $\mathbf{W}$ with probability $1-p$, as represented in Fig. 2.

## 4. Distributed detection method

### 4.1. Sensor network message design and dissemination

In this section, we propose a novel distributed failure detection algorithm and analyze its performance. The algorithm is based on a novel test design and message dissemination strategy in a distributed GT framework. The sensors iteratively create and disseminate messages in two-phases, denoted by $t_\mathrm{I}$ and $t_\mathrm{II}$. One algorithm round consists of these two phases. During the first phase $t_\mathrm{I}$, the sensors obtain messages that estimate the presence of defective sensors in their neighborhood. In the second phase $t_\mathrm{II}$, the sensors linearly combine messages and exchange them employing a gossip mechanism. They are illustrated in Fig. 3 and described below in more detail.

The first phase $t_\mathrm{I}$ in round $t$ represents the message construction process illustrated in Fig. 3(a). $L$ master sensors cluster the network into disjoint subsets $\mathcal{V}_l \subset \mathcal{V}$, $l = 1, \ldots, L$. Clustering is used to bound the search space of the decoder,
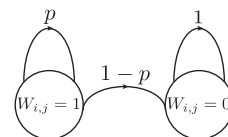


**Fig. 2.** Noise effect on binary symbols in the test message: non-zero values in the test matrix are flipped with probability $1-p$.

as explained in the following subsections. Measurements of neighbor sensors do not vary significantly when the sensors are not defective and when the signal under observation is smooth over the sensor field. The master sensors locally gather the readings or measurements from neighbor sensors. Each sensor randomly participates in the test with probability $q$, as given in Eq. (1). The master sensor estimates the presence of defective sensors within its neighborhood and then attributes a binary value $f(s_k) \in \mathbf{f}$ to each sensor in the neighborhood. The value $f(s_k) = 1$ denotes that the sensor $s_k$ is defective and $\mathcal{K}$ marks the set of defective sensors. Noise influence in the test is given as in Fig. 2. The test outcome at master node $l$ is finally computed as

$$g_l = \mathbf{W}_{l,:} \otimes \mathbf{f} = \begin{cases} 1 & s_k \in \mathcal{K}, \\ 0 & \text{otherwise}, \end{cases} \quad (3)$$

where the binary matrix operator $\otimes$ is composed of the bitwise OR $\odot$ and the bitwise addition $\oplus$ operators. The matrix $\mathbf{W} = [\mathbf{W}_{1,:}(t); \dots; \mathbf{W}_{B,:}(t)]$ that marks sensors that participated in tests is of size $B \times S$. The message formed by a master sensor $l$ during the phase $t_I$ consists of the outcome $g_l$ and the test participation identifier $\mathbf{W}_{l,:}$. The message $(g_l(t^-), \mathbf{W}_{l,:}(t^-))$ is sent to the neighbor sensors, which concludes the phase $t_I$.

During the phase $t_{II}$ the messages are disseminated within the network as illustrated in Fig. 3(b). Every sensor $s_i \in \{1, \dots, S\}$ requests the message formed at the previous round from its neighbor $s_j$, chosen uniformly at random, following a gossip mechanism with pull protocol. Next, $s_j$ responds to the message request by sending its message created in the previous round. This process is performed only once per round. The sensor $s_i$ further combines these messages as

$$g_i(t) \leftarrow g_i(t^-) \oplus g_j(t-1),$$
$$\mathbf{W}_{i,:}(t) \leftarrow \mathbf{W}_{i,:}(t^-) \oplus \mathbf{W}_{j,:}(t-1), \quad (4)$$

where $g_j(t-1)$ denotes the sensor outcome value of the neighbor $s_j$ at the previous round $(t-1)$. The vector $\mathbf{W}_{i,:}(t)$ represents the test indicator vector at the sensor $s_i$ in round $t$. Since the messages are created probabilistically, the message combination in the different rounds assures that an innovative message reaches sensors at every round with high

probability. A toy example of the dissemination phases is illustrated in Fig. 4. In this example the sensor $s_2$ at round $t$ pulls the message from the sensor $s_1$ and constructs a new message according to Eq. (4).

In a matrix form, the process of message formation and transmission in $B$ rounds at any sensor in the network is represented as

$$\mathbf{g} = \mathbf{W} \otimes \mathbf{f}. \quad (5)$$

This equation resembles to the outcome computation in the centralized GT case. However, in the distributed GT the tests represent linear combinations of test vectors that build disjunct matrix with high probability, as given in Eq. (1). We provide here an example to make a clear distinction between test matrices in proposed and centralized setup. We assume that an oracle has a direct access to the readings of the master nodes. A vector representation of tests performed at master nodes observed by an oracle at the round $t=i$ we denote with $\mathbf{C}_{i,:}$. Then, the matrix $\mathbf{C} = [\mathbf{C}_{1,:}; \mathbf{C}_{2,:}; \dots \mathbf{C}_{B,:}]$ represents the test matrix over $B$ collection rounds. Observe that the matrix $\mathbf{C}$ is by construction disjunct, while $\mathbf{W}$ is built on the Boolean addition of rows of $\mathbf{C}$ as in Eq. (4). The values in $\mathbf{W}$ thus depend on the random message propagation path, which is obviously not the case in the centralized GT algorithm. Note that, for an arbitrary network, the number of network rounds required for collecting a particular number of linearly independent tests varies and depends on the network topology, the number of master nodes $L$ and the test participation probability $q$. Once every sensor has gathered test messages, it independently computes the binary vector
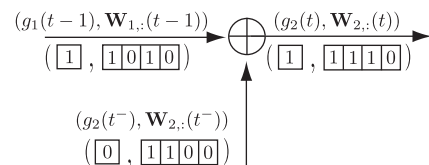


$$(g_1(t-1), \mathbf{W}_{1,:}(t-1)) \quad (g_2(t), \mathbf{W}_{2,:}(t))$$
$$(\boxed{1}, \boxed{1\,0\,1\,0}) \quad \bigoplus \quad (\boxed{1}, \boxed{1\,1\,1\,0})$$
$$(g_2(t^-), \mathbf{W}_{2,:}(t^-))$$
$$(\boxed{0}, \boxed{1\,1\,0\,0})$$

**Fig. 4.** The message formation at sensor $s_2$ in round $t$. We assume that $s_2$ pulls sensor $s_1$ to send its previous round values (round $t-1$). We assume that the sensor $s_3$ is defective $\mathbf{f} = [0010\dots]$. The outcome value and the test identifier vector are formed by bitwise OR.
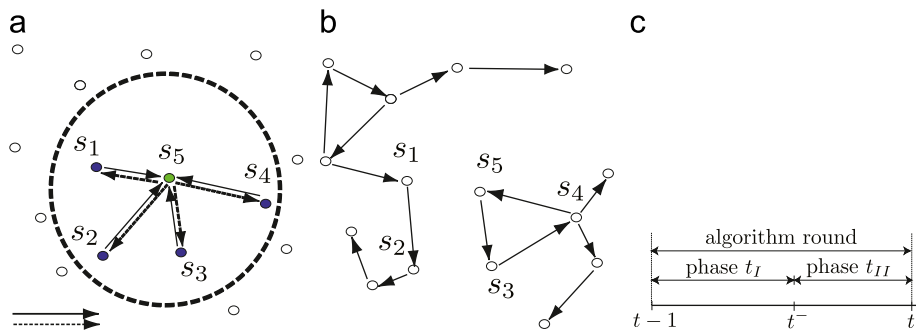


**Fig. 3.** Illustration of the message design and dissemination through the sensor network. (a) Message formation based on local sensor measurements: full and dashed arrows correspond to the steps of the message design, respectively. In the first step, the master sensor collects the sensor measurements from its neighbor sensors $\{s_1, \dots, s_4\}$ and forms the message $(g_l(t^-), \mathbf{W}_{l,:}(t^-))$. In the second step, the message is propagated from the master sensor to its neighbor sensors (phase $t_I$: Message design). (b) Message dissemination based on a gossip algorithm with pull protocol, where the sensors request the messages from their neighbors chosen uniformly at random (phase $t_{II}$: Message dissemination). (c) Rounds of communication in our iterative detection algorithm consist of the message design ($t_I$) and the message dissemination ($t_{II}$) phases (communication phases).

**f** that satisfies the tests in Eq. (5). This process is analyzed in more detail below.

### 4.2. Detection of one defective sensor in the network

We first analyze the case of a single defective sensor (case $K=1$) in the network and study the detection probability of our distributed algorithm. To recall, the distance decoder used for detection computes the Hamming distance between vectors. The element-wise distance is equal to one if element values differ, otherwise is zero. To avoid false alarms, the decoder threshold $\epsilon$ is set to a value higher than the expected number of noise-induced bit flips per columns in the disjunct matrix **C** [21]

$$\epsilon = (1+\delta)(1-p)qB, \tag{6}$$

where $\delta > 0$ is a small constant and $B$ is the number of rows in **C**. Columns of **C** have in average $qB$ non-zero elements and non-zero elements are flipped with probability $(1-p)$, so the expected number of flips per column is

$$\mu = (1-p)qB. \tag{7}$$

Recall that the matrix **C** is by construction a disjunct matrix. The detection problem is resolved for tests that form a disjunct test matrix [21, Proposition 2]. However, the messages available at sensors in the network form a test matrix that is obtained by linear combinations of disjunct matrix rows and not disjunct matrix rows itself. Nevertheless, we show below that the distance decoder detects a defective sensor with high probability under certain conditions.

The formal propositions for detection with high probability are given below. Proposition 1 and the first part of Proposition 2 that refer to designing a $(K,\epsilon)$-disjunct matrix with high probability are similar to Proposition 2 of [21], which is however derived for centralized detection. They represent the starting point of the analysis of our distributed algorithm and we include them here because they were not used previously in the context of distributed detection in sensor networks. First we show that for networks with a single master node the proposed algorithm designs a $(K,\epsilon)$-disjunct matrix **C** during the phase $t_I$. Next we show that linear combinations of rows of **C** preserve distances between the test outcome and the column of the defective sensor in the test matrix. We then build on the first two propositions to analyze the number of messages needed for the distributed detection of a single defective sensor, which is given in Proposition 5.

Below we show that networks with a single master node ($L=1$) and probabilistic message design in $t_I$ build a $(K,\epsilon)$-disjunct matrix **C** with high probability. This case boils down to the centralized collection of data given in Proposition 2 in [21].

**Proposition 1.** *For a single-cluster network, the message design over the phase $t_I$ of our proposed method builds a $(K,\epsilon)$-disjunct matrix **C** with high probability for an arbitrary $K$ and $\epsilon$ defined as in Eq. (6).*

**Proof.** We show that the probability that the number of rows with a good disjunctness property $G$ of **C** is smaller than $\epsilon$ and we follow the development proposed in [21]. The sensor participation probability $q$ in a test is defined

as in Eq. (1). A row of the matrix $\mathbf{C}_{i,:}$ is considered to have a good disjunctness property if a single symbol "1" occurs, while the rest $K-1$ values are equal to zero. The probability of such an event is equal to $\mu_1 = q(1-q)^{K-1}$. The random variable that marks the total number of rows with such a property is denoted with $G$. The distribution of $G$ is binomial with a mean value $\mu_2 = \mu_1 B$. We show that the probability of having less than $\epsilon$ rows with a good disjunctness property is small under the assumption that $\epsilon < \mu_2$. We limit this probability by a Chernoff bound

$$P(G<\epsilon) \le e^{-(1/2)(\mu_2-\epsilon)^2/\mu_2} = e^{-qB[(1-q)^{K-1}-(1-p)(1+\delta)]^2/2(1-q)^{K-1}}. \tag{8}$$

We mark the exponential term as $\gamma = [(1-q)^{K-1}-(1-p)(1+\delta)]^2/2(1-q)^{K-1}$. Since $2 < e < 3$, $2^{-m} \ge e^{-m} \ge 3^{-m}$ and constant $m \ge 0$, $\gamma$ is bounded. For the parameter choice in [21] $(\delta,m)=(p/2,p/8)$, the value $\gamma = \mathcal{O}(p^3)$. Therefore this probability can be designed to be arbitrarily small: $P(G<\epsilon) \le e^{-B\gamma/K} = e^{-\mathcal{O}(Bp^3/K)}$. □

Next, we show that linear combinations of rows of **C** for network with one master node preserve the Hamming distance only between the column of matrix $\mathbf{W}_{:,k}$ that corresponds to the defective sensor $s_k$ and the outcome vector **g**.

**Proposition 2.** *Let **C** be the $(K,\epsilon)$-disjunct matrix created over $B$ consecutive rounds in a single-cluster network during the phase $t_I$. Linear combinations of messages generated during the phase $t_{II}$, performed as in Eq. (4), preserve the Hamming distance between the column of obtained matrix $\mathbf{W}_{:,k}$ that corresponds to the defective sensor $s_k$ and the outcome vector **g**.*

**Proof.** We first analyze the number of value flips which leads to a decoding failure for $(K,\epsilon)$-disjunct matrices, following a development similar to [21]. Next, we prove that linear combinations of rows in such matrices preserve vector distances between the outcome vector and the column of **W** that corresponds to the defective sensor.

A decoding failure with a distance decoder occurs in a $(K,\epsilon)$-disjunct matrix when the number of flips of column elements of **C** is higher than $\epsilon$. The probability of an occurrence of a single flip is equal to $\mu_3 = q(1-p)$. Let $F$ denotes the number of flips in the columns of the matrix. Hence, the expected number of flips per column is given in Eq. (7). We want to compute the lower bounds for the event that more than $(1+\delta)\mu$ flips occurred in the column of the matrix, for $\delta > 0$. Applying the Markov inequality $P(F \ge (1+\delta)\mu) \le \inf_{d>0} \prod_{i=1}^{S} E[e^{dF_i}]/e^{d(1+\delta)\mu}$ for a constant $d > 0$ and plugging the probability of the single flip event:

$$P(F_i) = \begin{cases} 1 & \text{with probability } (1-p)q, \\ 0 & \text{with probability } 1-(1-p)q, \end{cases} \tag{9}$$

to the expectation term of the previous equation leads to

$$P(F \ge (1+\delta)\mu) \le \inf_{d>0} \frac{\prod_{i=1}^{S}[(1-p)q(e^d-1)+1]}{e^{d(1+\delta)\mu}}. \tag{10}$$

Using simple computations and setting $x = (1-p)q(e^d - 1)$, $d = \log(1+\delta)$ and $1+x < e^x$ we finally obtain

$$P(F \geq (1+\delta)\mu) \leq e^{-\mu\delta^2/(2+\delta)}. \tag{11}$$

The outcome value $\mathbf{g}$ depends on the presence of a defective sensor $s_k$ in the test. We prove next that the distance between $\mathbf{g}$ and the $k$-th column $\mathbf{W}_{:,k}$ does not increase more than $\epsilon$ during $t_{\text{II}}$, while this is not true for the rest of the columns. When sensor $s_j$ sends its message to sensor $s_i$ during the round $t$, we have

$$dist(g_i(t), W_{i,k}(t)) = dist(g_i(t^-) \oplus g_j(t-1), W_{i,k}(t^-) \oplus W_{j,k}(t-1))$$
$$= dist(g_i(t^-), W_{i,k}(t^-)) \oplus dist(g_j(t-1), W_{j,k}(t-1)), \tag{12}$$

where the first equality results from Eq. (4). The second equality follows directly from the fact that the values of $\mathbf{g}(t^-)$ and the columns $\mathbf{W}_{:,k}(t^-)$ are identical for the defective sensor due to Eq. (5). Since these two columns initially may differ at $\epsilon$ positions due to noise flips, the overall distance between the vectors $\mathbf{g}(t^-)$ and $\mathbf{W}_{:,k}(t^-)$ is at maximum $\epsilon$ defined by Eq. (6). □

We now consider networks with $L$ master sensors and a hypothetical centralized data collection. We assume that $L$ master nodes cluster the sensor network in disjoint subsets, where every sensor belongs to exactly one cluster. The master nodes perform message design over the rounds $t_I$ as proposed by our algorithm. We show that the tests gathered from the $L$ different clusters build a disjunct matrix, where each cluster relates a $(K,\epsilon)$-disjunct matrix.

**Proposition 3.** *The diagonal matrix* $\mathbf{C} = diag(\mathbf{C}_1, \ldots, \mathbf{C}_L)$ *obtained from* $(K,\epsilon)$-*disjunct matrices* $\mathcal{C} = \{\mathbf{C}_i\}_{i=1}^L$ *is at least* $(K,\epsilon)$-*disjunct.*

**Proof.** The proof follows directly from the disjunctness property of matrices in $\mathcal{C}$ (Propositions 1 and 2). □

We analyze now the influence of message gathering over successive rounds of our detection algorithm. Uniform gathering of linearly combined messages at $L$ clusters by a hypothetical centralized decoder enables the detection of the defective sensor with high probability when the number of received messages is sufficient.

**Proposition 4.** *When the* $(K,\epsilon_i)$-*disjunct matrices* $\mathcal{C} = \{\mathbf{C}_i\}_{i=1}^L$ *are linearly combined as in Eq. (4), where* $\epsilon = \sum_{i=1}^L \epsilon_i$ *and* $q = \sum_{i=1}^L q_i$, *the resulting test matrix permits detection by a distance decoder with high probability as long as* $B \geq \mathcal{O}(K \log(S)/p^3)$ *messages collected from randomly chosen clusters build this matrix.*

**Proof.** We first show that a diagonal matrix constructed from $(K,\epsilon_i)$-disjunct matrices of the set $\mathcal{C}$ is $(K,\epsilon)$-disjunct. Next, we recall Proposition 2 and finally, we show that the $B$ measurements assure a good disjunct property of cluster matrices. Let the number of rows for all matrices be $B = \mathcal{O}(K \log(S)/p^3)$. The parameters $\epsilon$ and $\epsilon_i$ are defined in Eq. (6) and $\epsilon = \sum_{i=1}^L \epsilon_i = (1+\delta)(1-p)Bq$, so the diagonal matrix built out of $(K,\epsilon_i)$ matrices is $(K,\epsilon)$ disjunct. The next part of the proof follows from Proposition 2 which states that a matrix whose rows are formed by linear

combinations of rows of $(K,\epsilon)$-disjunct matrix permits detection with a distance decoder. Finally, we prove that for a given matrix $\mathbf{C}$ the disjunct property holds if at least $B$ messages are available. For this purpose, we follow a development similar to [21] and consider that the number of sensors in clusters is equal to the total number of sensors $n = S$. The probability bound given in Proposition 1 should hold for all possible choices of a fixed set of $T$ out of $S$ columns: $\bigcup_T P(G \leq \epsilon) \leq Se^{-B q \gamma}$. This probability can be arbitrarily small, e.g., in case $B \geq K \log S / m\gamma = \mathcal{O}(K \log S/p^3)$. Further on, the condition in Eq. (11), which gives the probability bound that the number of flips in any $K$ out of $T$ columns exceeds a threshold value $\epsilon$ is also bounded. It reads

$$\bigcup_K P(F \geq (1+\delta)\mu) \leq Ke^{(-\delta^2/(2+\delta))\mu} = Ke^{(-\delta^2/(2+\delta)p^3)(1-p)qK \log(S)},$$

where the last equality is obtained from Eq. (7). This probability is small for the sufficiently large value of $B = \mathcal{O}(K \log(S)/p^3)$. □

We now analyze the proposed distributed algorithm and consider the detection requirements for every sensor in the network. We show that the test messages collected by the sensors during the transmission rounds enable failure detection by the distance decoder with high probability if the number of messages is sufficient, where the decoder operations are performed locally at sensors.

**Proposition 5.** *We assume that* $L$ *master sensors partition the sensor network into disjunct parts. Test realizations within a cluster form test vectors. Over the rounds, these vectors create* $(K,\epsilon)$-*disjunct matrices* $\mathcal{C} = \{\mathbf{C}_i\}_{i=1}^L$ *whose elements*

$$c_{i,j} = \begin{cases} 1 & \text{with probability } q_i = \alpha_i, \\ 0 & \text{otherwise}, \end{cases} \tag{13}$$

*where* $q = \sum_{i=1}^L q_i$. *Messages* $(\mathbf{g}_i, \mathbf{W}_{i,:})$ *arrive at all the sensors in the network in our proposed algorithm, as described in the previous section. If the above assumptions hold and if the number of linearly independent messages received per cluster at every sensor in the network is at least* $B/L$, *where* $B \geq \mathcal{O}(K \log(S)/p^3)$, *the probability that sensors fail to detect the defective sensor by the distance decoder tends to zero as* $S \to \infty$.

**Proof.** The message collection method does not influence the decoder performance, since the number of per-cluster measurements is sufficient for decoding with high probability. Therefore, the proof follows from the proof of Proposition 4. □

### 4.3. Detection of multiple defective sensors in the network

We analyze now the distributed detection of multiple defective sensors where $K \ll S$ holds. We propose here to slightly modify our distributed algorithm and to limit the decoder search space to be able to apply the Hamming distance decoder. The protocol modification and the adaptation of the distance decoder are described below. We assume that sensors completely differentiate between sensors in the network that belong to particular clusters

and that at most one defective sensor is located in a given cluster. This knowledge limits the size of the decoder search space.

The proposed protocol is first modified as follows to deal with multiple defective sensors. A decoder error occurs when two or more messages with positive test outcomes are combined together during the phase $t_{II}$, since the distance preserving property defined in Eq. (12) is not guaranteed in this case. Since the number of defective sensors is very small compared to the total number of sensors, this event however occurs rarely. We explain the protocol modification with a simple example. Let the sensor $s_i$ pull the message from the sensor $s_j$, where both sensor test outcomes have non-zero values. Instead of combining the messages as in Eq. (4), we simply buffer the new message of $s_i$ and consider the message from $s_j$ at previous round as the final outcome of the phase $t$:

$$g_i(t) = g_j(t-1),$$
$$\mathbf{W}_{i,:}(t) = \mathbf{W}_{j,:}(t-1). \tag{14}$$

At the first subsequent round $\tau \geq t+1$ of our distributed algorithm where both messages $g_i(\tau)$ and $g_j(\tau-1)$ have non-zero values, $g_i(\tau)$ is replaced by the message buffered in the node $s_i$. The rest of the protocol remains unchanged.

The decoding proceeds in two main steps. First, the appropriate unions of test matrix columns are created to form the search set space and second, the Hamming distance between the test outcome vector and the vectors of the search set is computed. The minimum Hamming distance indicates the solution of the detection problem. The outcomes $\mathbf{g} = [\mathbf{g}_0 \ \mathbf{g}_1]^T$ collected at some sensor are divided into two sets, i.e., the negative and positive outcome vectors $\mathbf{g}_0$ and $\mathbf{g}_1$, respectively. Subsequently, the rows of the test matrix $\mathbf{W}$ form two sub-matrices $\mathbf{W}_0$ and $\mathbf{W}_1$ and Eq. (5) is rewritten as

$$\begin{bmatrix} \mathbf{g}_0 \\ \mathbf{g}_1 \end{bmatrix} = \begin{bmatrix} \mathbf{W}_0 & 0 \\ 0 & \mathbf{W}_1 \end{bmatrix} \begin{bmatrix} \mathbf{f}_0 \\ \mathbf{f}_1 \end{bmatrix}. \tag{15}$$

We eliminate non-defective sensors from $\mathbf{W}_1$ using the knowledge from $\mathbf{W}_0$ and obtain $\mathbf{W}_1'$. To remind, master nodes partition sensors into clusters. The columns of interest are those columns of $\mathbf{W}_1'$ which contain at least one non-zero value, since they mark the participation of the potential defective sensors in tests. We build the total search space $\mathcal{U}$ by taking into account the total or partial knowledge of sensor cluster partition.

If the full cluster partition is known, the columns of the sensors that belong to the same cluster $l$ are grouped together in a set $\mathcal{H}_l$, $l = \{1, \ldots, L\}$, where $L$ is the total number of clusters (master nodes). Given a partial knowledge about the sensor cluster partition, the cardinality of the set $|\mathcal{H}| = |\{\mathcal{H}_l\}_{l=1}^{L^*}|$ is $L^* > L$. The search space $\mathcal{U}$ consists of vectors that are obtained by element-wise OR addition of up to $K$ vectors that are picked from different sets $\mathcal{H}_l$, because of the assumption that at most one defective sensor exists in each cluster. For instance, let the number of defective sensors and clusters be $(K,L) = (2,2)$. Let $\mathcal{H}_1$ contain $h_1$ and $\mathcal{H}_2$ contain $h_2$ columns. Then the search space size has in total $|\mathcal{U}| = h_1 h_2 + h_1 + h_2$

elements, where $h_1 h_2 = \binom{h_1}{1} \cdot \binom{h_2}{1}$ denotes the number of unions of $K = 2$ columns and single column subsets are chosen in $h_1 + h_2$ ways. Distance decoding is performed between $\mathbf{g}_1$ and elements of the set $\mathcal{U}$, starting from the vectors that are created as unions of $K$ columns towards the smaller number of column unions. If no solution exists for a particular value of $K$, we perform the decoding for vectors built from $K-1$ column unions of $\mathcal{H}_i$. If no unique solution is found, we encounter a decoding failure.

We now analyze the number of messages that are required for detection of multiple defective sensors with high probability.

**Proposition 6.** *Under the assumption that at most one defective sensor is present in each cluster, that the number of available linearly independent messages at all sensors is at least $B/L$ per cluster with $B \geq \mathcal{O}(K \log(S)/p^3)$ and that sensors know membership identifiers of all the clusters in the network, the distance decoder detects defective sensors at all sensors in the network with high probability.*

**Proof.** To recall, the transmission protocol ensures that the assumptions imposed by Proposition 5 hold for one defective sensor. Then, due to the assumption that at most one defective sensor is present in one cluster and that there is at most one defective sensor active in the test, we form the set of solutions for the multiple defective case, which has a unique solution. Distance decoder between the outcome vector and a limited set of vectors that forms a full search space can therefore find the appropriate solution. In other words, this procedure is identical to per-cluster decoding, where each cluster has at most one defective element, so Proposition 5 can be applied. □

**Proposition 7.** *Under the assumption that at most one defective sensor is present in each cluster, that the number of available linearly independent messages at all sensors in the network is at least $B/L$ per cluster, where $B \geq \mathcal{O}(K \log(S)/p^3)$ and sensors know the partial set of identifiers of the clusters in the network, the distance decoder detects defective sensors at all sensors in the network with high probability.*

**Proof.** The search space $\mathcal{U}$ created in this case is larger, but it contains the solution. Now the proof is identical to that in the previous proposition. □

The assumption that at most one defective sensor occurrence per cluster is reasonable when $K \ll S$ and it is easy to bound the probability that at least two defective sensors occur within any cluster with the particular combination of $(S,L,K)$ parameters. We drop this analysis due to space limitations. Details can be found in our Technical Report [22, Section III.C].

## 5. Performance evaluation

### 5.1. Setup

In this section, we investigate the performance of our distributed detection method denoted as GP in various scenarios. We first examine the influence of different network parameters on the rate of message dissemination. Next, we examine the decoding probability for

defective sensor(s) detection. The number of system rounds required to collect the necessary number of messages for accurate decoding varies with the network topology. The simulations are performed for fully connected, $k$-connected and irregular graphs. Finally, we discuss the number of required linearly independent measurements for successful detection and compare it with the theoretical value.

We also analyze the performance of several alternative schemes, namely a random walk method that employs a gossip mechanism with pull protocol (RWGP) and a classical random walk (RW) detection. A random walk determines the path of successive random dissemination message exchanges between neighbor sensors. In the RWGP method, the random walk is initiated at $L$ sensors (equivalent to the master sensors in the GP method) and terminates after a pre-determined number of rounds. The sensors create messages from the sensor measurements collected along the random walk path. These messages are transmitted with the gossip algorithm that uses a pull protocol. Note that, for identical choice of the sensors over rounds, RWGP and GP are identical. The RW method initiates the raw (uncompressed) measurements collection in $L$ random sensors and completes it in a given number of rounds. Every sensor that lays along the random walk path stores the values of all sensors along the transmission path. When all the sensors receive all data, the process terminates.

The proposed GP algorithm is also compared with a store-and-forward (SF) and a greedy store-and-forward (GSF) method that employs pull protocol. Both algorithms disseminate raw sensor measurements. For the SF method, upon receiving a message request, a node responds by forwarding randomly chosen messages from the available set of messages. In GSF, each sensor randomly requests the innovative measurements in a greedy manner from its randomly chosen neighbor sensor. This procedure involves the additional message exchange among sensors in every round.

We analyze the performance of these algorithms in fully connected, $k$-regular and irregular networks. We randomly position sensors in a unit square area to obtain irregular topologies. Sensors that lay within a certain radius may exchange messages directly. In each case, we build 10 different network realizations and for each such realization we perform 100 independent simulations. The results of all the sensors in the network are averaged over all the simulations. Detection failure is encountered when at least one sensor value in the network is decoded wrongly.

### 5.2. Detection performance

We first consider the case of a single defective sensor ($K=1$). The detection probability and the average rank evolution over rounds are examined for fully connected (FG) and $k$-connected regular networks (RG) with sensors degree $k \in \{6,16\}$. Here, the network consists of $S=20$ sensors. From Fig. 5, we see that networks with higher number of connections achieve faster dissemination of the innovative messages. We also note that high connectivity value $k$ is beneficial, but it cannot drive by itself the performance of our detection scheme. It should be combined with the appropriate choice of network parameters, discussed in more detail in Technical Report [22]. For example, RM master sensor selection for $k=16$ achieves better detection performance, compared to that of fully connected graphs.

In Fig. 6, we illustrate the detection probability for random graphs (100 simulations per different graph) with $S=70$, $K=1$ defective sensor and minimum sensors' degree $k \geq 3$. We observe that random graphs require more rounds in average for successful detection, as expected. Also, we observe that the detection performance decreases because of the limited message diversity (smaller probability of receiving innovative messages) and the lower connectivity.

We then consider the case of multiple defective sensors. In Fig. 7, we present results for the cases with $K=2$ defective sensors in networks with 70 sensors. The results are given in terms of the average detection probability (for 10 different network realizations, we perform 100 independent simulations and calculate their average) over dissemination rounds, for both fully and irregularly connected graphs. The master sensors are selected deterministically (DM) due to the decoder design for multiple
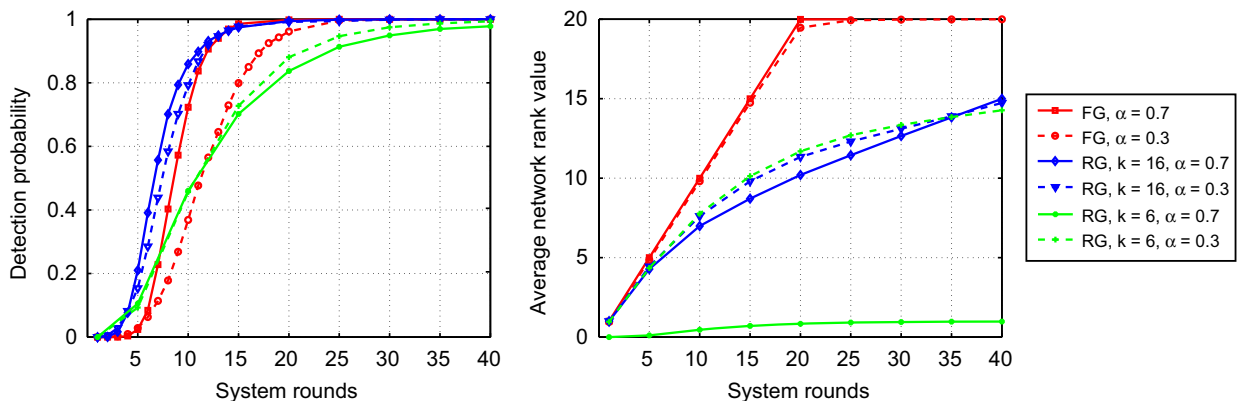


**Fig. 5.** Simulation results for fully connected (FG), $k=16$-regular connected (RG, $k=16$) and $k=6$-connected graphs (RG, $k=6$) and sensor participation constant $\alpha = qK$ with $S=20$ sensors, $K=1$ and a random selection (RM) of $L=5$ master sensors. (Left) Probability of defective sensor detection. (Right) Average rank of messages received per sensor.
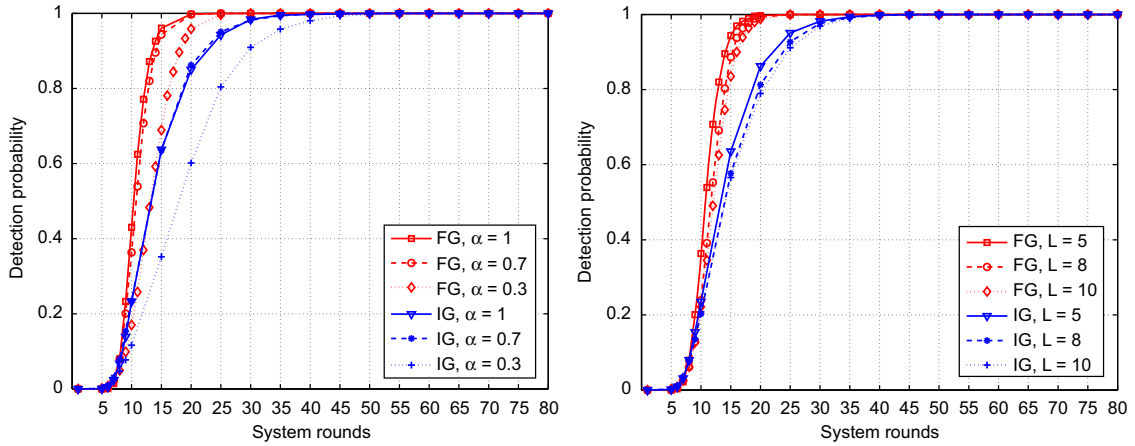
**Fig. 6.** Probability of defective sensor detection; simulation results for irregular graphs ($k > 3$) and random selection (RM) of $S=70$ sensors, $K=1$. (Left) $L=5$ master sensors. (Right) Sensor participation constant $\alpha = qK = 0.7$.
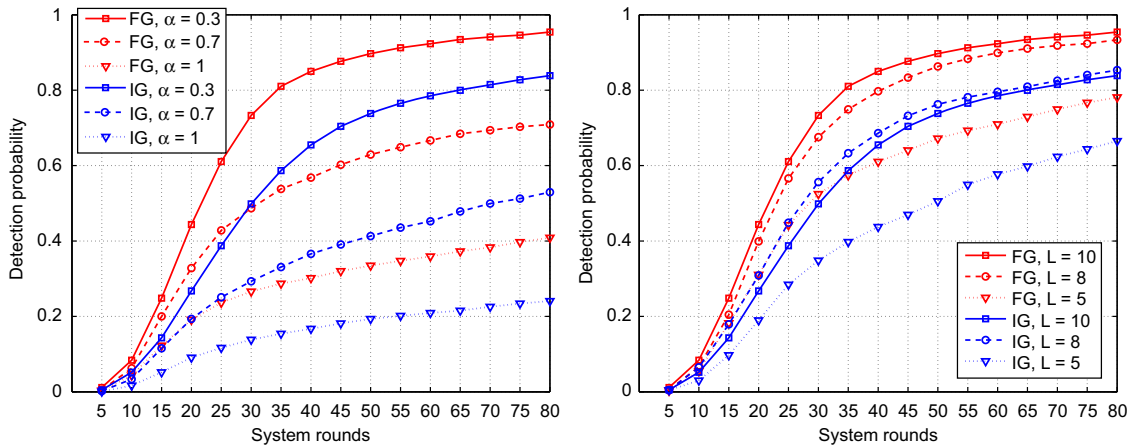


**Fig. 7.** Probability of defective sensor detection for fully connected (FG) and irregular graphs (IG), $d > 3$ with $S=70$ sensors, $K=2$ and deterministic selection (DM) of master sensors. (Left) Fixed number of clusters $L=10$. (Right) Fixed sensor participation constant $\alpha = qK = 0.3$.

defective sensors identification. We focus on the evolution of the decoding probability and the average number of messages collected over rounds. From the evaluation it is clear that the detection performance is reasonable when the selected parameter values $(L, \alpha)$ favor diverse message generation.

In [21], a centralized system has been proposed, which can be considered as dual to fully connected networks with centralized tests (single master sensor that covers all the network). For comparison reasons, we compute the required number of measurements for networks with $S \in \{20, 70\}$ sensors and parameters $(q, pf_1, pf_2) = ((0.15-0.3), 0.01, 0.01)$. The results are reported in Table 1. We observe that the worst case analysis leads to a higher number of dissemination rounds than the real ones. However, these values decrease relatively to the growth of number of sensors in the network. Simulations show that in practice the required measurements are significantly fewer.

Detection probability comparison of the proposed method with several random walk detection methods is

**Table 1**
The theoretical measurement requirements for networks with $S$ sensors.

|  | $S=20$ |  | $S=70$ |  |
|---|---|---|---|---|
|  | $K=1$ | $K=2$ | $K=1$ | $K=2$ |
| $p \in (0.9-1)$ | 130 | (115–244) | (174–217) | (125–284) |

illustrated in Fig. 8 for $S=70$ sensors. The proposed scheme outperforms all other comparison methods. Note that the number of required rounds in RWGP scheme for a high probability detection is large compared to the other schemes, while RW needs higher communication overhead for dissemination due to the transmission of raw sensor measurements. Average rank values over the network rounds are illustrated in Fig. 9. We observe that for the fixed detection probability $p=0.9$ for the network with $S=70$ sensors the average number of system rounds required for the proposed method is approximately $\{17, 20\}$. The number of system rounds required by the
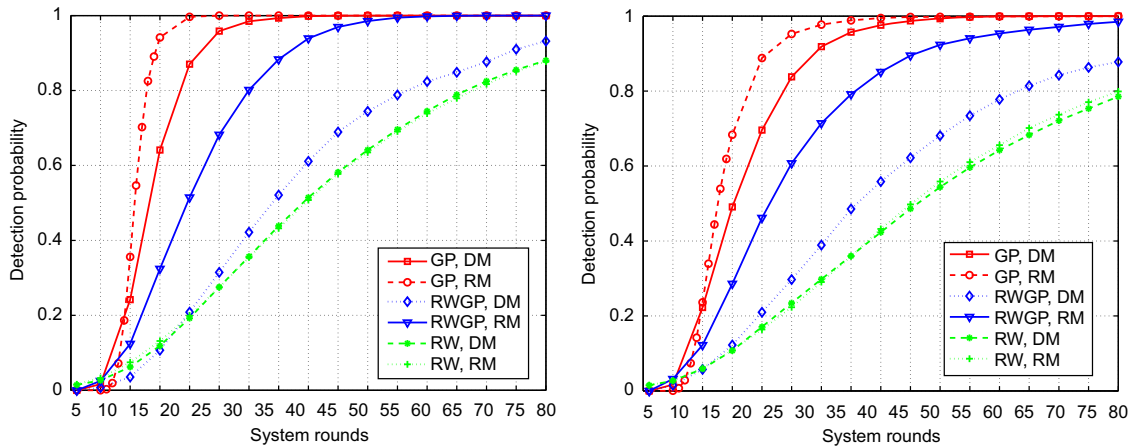
**Fig. 8.** Detection performance for networks with $S=70$ sensors and $L=5$ master sensors. Abbreviations: GP: proposed method, RWGP: random walk gossip dissemination algorithm with pull protocol, RW: random walk in the network initiated at $L$ sensors. (Left) Fully connected sensor network. (Right) Irregular sensor network.
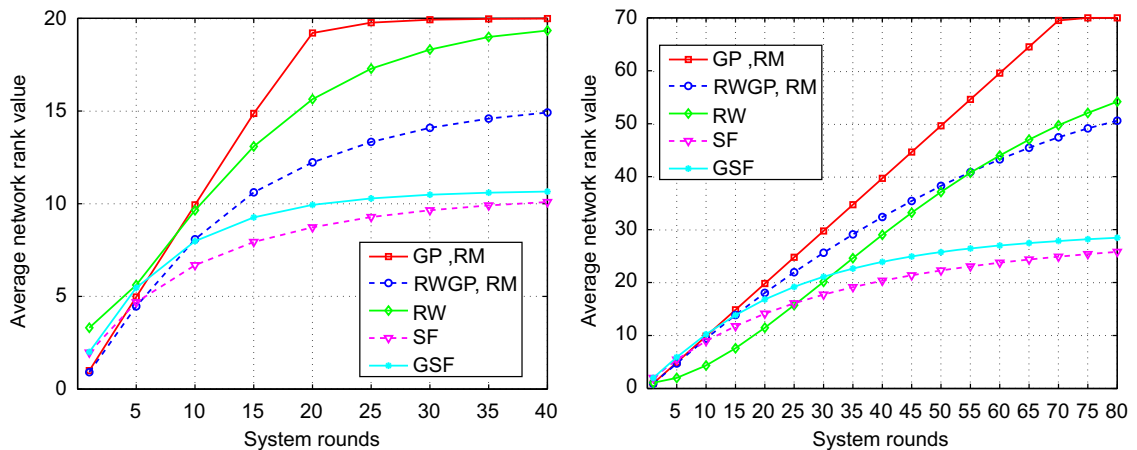


**Fig. 9.** Average rank value for irregular sensor networks with $L=5$ master sensors: (Left) $S=20$ sensors. (Right) $S=70$ sensors. Abbreviations: GP: proposed method, RWGP: random walk gossip dissemination algorithm with pull protocol, RW: random walk in the network initiated at $L$ sensors, SF: pull store-and-forward algorithm with a random message transmission, GSF: pull store-and-forward algorithm with a greedy message transmission.

other algorithms to reach the same probability of performance is higher.

We note here that the class of defective sensor signals which can be successfully detected is smaller in GP than in the centralized GT method because of the additional assumption that the defective sensors are scattered in the network. Those sparse signals whose non-zero values are localized may be only partially recovered with the proposed approach due to the partitioning into clusters. However, the cardinality of the class of locally distributed sparse signals is much smaller than the number of realizations of scattered sparse signals that this algorithm detects. Also, we highlight that the main advantages of the proposed distributed algorithm compared to the centralized one are reflected in lower communication costs and higher robustness. We finally note that any sensor can act as a decoder, which is useful in scenarios with one or more mobile decoders.

We do not attempt to optimize the parameter values in this work but we briefly discuss the influence of the

parameter choice in the paper. Details are provided in Technical Report [22], where we discuss the parameter values which increase the novelty probability of messages for different master node selection cases. The number of message exchanges in the network depends on the network connectivity and the number of master nodes $L$. This parameter can be optimized for each mode of the master node selection (deterministic/probabilistic) when the topology and sensor positions are known. These details are omitted because of the space limitations.

### 5.3. Communication overhead

For the sake of completeness, we analyze the communication costs of the proposed gossiping protocol and compare it with all other schemes under comparison. Let $R_d$ and $I_d$ denote the number of bits needed for raw measurements transmission and sensor identifier, respectively. Recall that the tuple $(S,L,L_n,n,\tau)$ stands for the number of sensors in the network, the number of master sensors (clusters), the

number of neighbors that each master is connected with, the average number of sensors per cluster ($n=S/L$) and the total number of transmission rounds, respectively.

During the first phase of GP, the master sensors receive raw measurements from their neighbors. Thus, $L_n \cdot R_d$ bits are used for communicating these values. Further, the master sensors create binary messages and send them to their neighbors. Every neighbor requires knowledge about the test sensor identifiers, thus the cost is $I_d \cdot \lceil q(L+L_n) \rceil$ bits, plus an additional bit in each message for sending the outcome result. Hence, the overall bit consumption is $L_n R_d + L_n(I_d \lceil q(L+L_n) \rceil) + 1)$. $(S+1)$ bits out of required $S(S+1)$ bits in the message exchange phase are reserved for the test outcome and the test matrix row. Note that this analysis includes the full vector size and it can be further compressed. The overall number of transmitted bits over $\tau$ rounds is

$$n_{GP}^b = \tau[L_n\{R_d + I_d \lceil q(L+L_n) \rceil + 1\} + S(1+S)]. \tag{16}$$

We compare the GP communication costs with the one of RWGP that also has two phases. The first phase represents the random walk message collection, while the second is equivalent to the GP algorithm. Note that the identical decoding performance of RWGP and GP occurs in a special case, when both of these algorithms collect exactly the same data. However, if in RWGP data collection from some sensors occurs multiple times (more probable in irregular networks with a smaller connectivity degree), it performs worse than GP. In typical simulations, a random walk of RWGP terminates after $n$-th transmission round, where $n$ is the number of elements per cluster in GP. RWGP transmits raw measurements, which results in $(1+n)R_d/2$ bits. Therefore, the communication cost for RWGP is given by

$$n_{RWGP}^b = \tau \left[ \frac{(n+1)R_d}{2} L + S(1+S) \right]. \tag{17}$$

The bit transmission requirements for the *RW* algorithm is equivalent to that of the first step of RWGP, since it also transmits raw data. The detection is performed at nodes by comparison of known sensor values at that moment, excluding the message design step. The number of transmitted bits is equal to: $n_{RW}^b = \tau((n+1)R_d/2)L$. Recall that one requires $\log S$ transmissions for a message dissemination to all the nodes in a fully connected graph. Therefore, SF algorithm requires in total $n_{SF}^b = \tau R_d \log S$ bits.

The comparison between the proposed method and all other schemes regarding the bits spent on communication is illustrated in Fig. 10 for a fully connected graph. Note that the proposed algorithm in this setup requires only $t=15$ rounds for detection, but it consumes approximately three times more communication overhead compared to that of RWGP algorithm. However, due to the specific collection approach (hops), the duration of one transmission round of RWGP lasts 10 times longer than that of the proposed algorithm. From the figure we can observe that the RW algorithm has very small communication overhead. However, it requires significantly higher number of rounds ($S \log S \approx 130$ rounds) compared to the detection time of the proposed GP algorithm. Overall, the proposed GP scheme is able to compete with the other schemes in terms of used bits until detection.

## 6. Conclusion

In this work, we have addressed the problem of distributed failure detection in sensor networks. We have proposed a novel distributed algorithm that is able to detect a small number of defective sensors in networks. We have designed a probabilistic message propagation algorithm that allows the use of a simple and efficient distance decoder at sensors. The transmitted messages are formed from local sensor observations and they are communicated using a gossip algorithm. For the worst case scenario we have derived the lower bound on the required number of linearly independent messages per cluster that sensors need to collect to ensure detection of
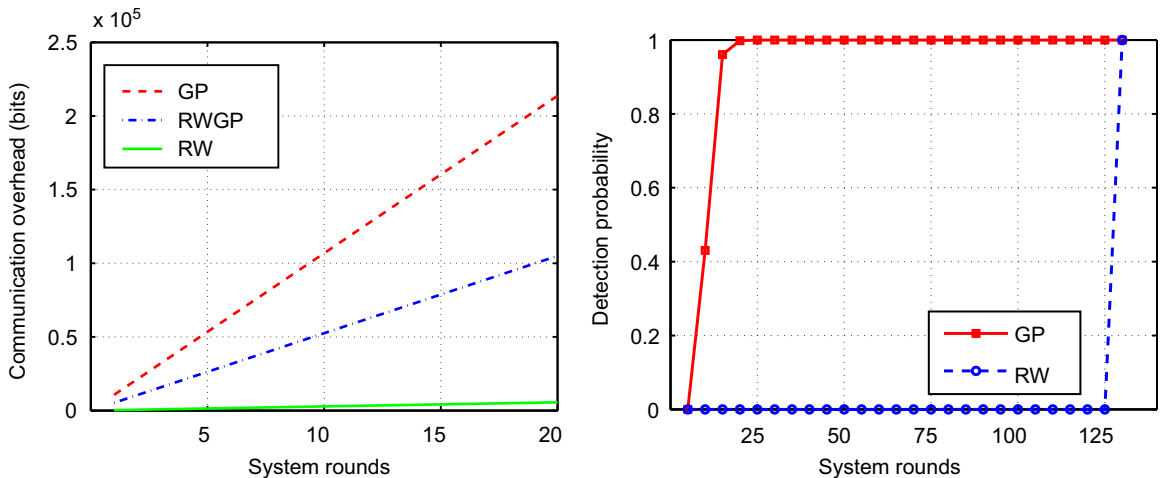


**Fig. 10.** (Left) Comparison of the communication overhead for several algorithms, for the following parameter values: $(S, L, L_n, \alpha, R_d, I_d, \tau) = (70, 5, 50, 0.7, 7, 7, 80)$. Graph is fully connected. Abbreviations: GP: proposed method, RWGP: random walk rounds with gossip algorithm and pull protocol dissemination, RW: random walk in the network initiated at $L$ sensors. (Right) Comparison of detection vs. number of rounds of the distributed detection scheme.

one defective sensor with high probability. We have shown experimentally that this number is quite smaller in practice, even for the small size networks, which confirms the validity of the theoretical bound. The experimental results have shown that the proposed method outperforms other detection schemes in terms of successful detection probability. The convergence rate is very fast, which largely compensates for the higher communication overhead.

# References

[1] R. Dorfman, The detection of defective members of large populations, Annals of Mathematical Statistics 14 (1943) 436–440.

[2] A. Dimakis, S. Kar, J. Moura, M. Rabbat, A. Scaglione, Gossip algorithms for distributed signal processing, Proceedings of the IEEE Transactions on Information Theory, 2010, pp. 1847–1864.

[3] A. Demers, D. Greene, C. Hauser, W. Irish, J. Larson, S. Shenker, H. Sturgis, D. Swinehart, D. Terry, Epidemic algorithms for replicated database maintenance, Proceedings of the 6th Annual ACM Symposium on Principles of Distributed Computing, 1987, pp. 1–12.

[4] P.K. Varshney, Distributed Detection and Data Fusion, 1st edition, Springer-Verlag New York, Inc., 1996.

[5] J.N. Tsitsiklis, Decentralized detection, Proceedings of Advanced Statistical Signal Processing, Signal Detection, vol. 2, 1993, pp. 297–344.

[6] Q. Tian, E.J. Coyle, Optimal distributed detection in clustered wireless sensor networks, IEEE Transactions on Signal Processing 55 (2007) 3892–3904.

[7] R. Viswanathan, P.K. Varshney, Distributed detection with multiple sensors: part I—fundamentals, Proceedings of the IEEE 85 (1997) 54–63.

[8] R.S. Blum, S.A. Kassam, H.V. Poor, Distributed detection with multiple sensors: part II—advanced topics, Proceedings of the IEEE 85 (1997) 64–79.

[9] M. Young, R. Boutaba, Overcoming adversaries in sensor networks: a survey of theoretical models and algorithmic approaches for tolerating malicious interference, IEEE Communications Surveys and Tutorials 13 (2011) 617–641.

[10] D.-Z. Du, F.K. Hwang, Combinatorial Group Testing and its Applications, 2nd edition, World Scientific, 2000.

[11] H.-B. Chen, F.K. Hwang, A survey on nonadaptive group testing algorithms through the angle of decoding, Journal of Combinatorial Optimization 15 (2008) 49–59.

[12] W. Dai, O. Milenkovic, Weighted superimposed codes and constrained integer compressed sensing, IEEE Transactions on Information Theory 55 (2009) 2215–2229.

[13] A.D. Bonis, U. Vaccaro, Constructions of generalized superimposed codes with applications to group testing and conflict resolution in multiple access channels, Theoretical Computer Science 306 (2003) 223–243.

[14] M.T. Thai, Y. Xuan, I. Shin, T. Znati, On detection of malicious users using group testing techniques, International Conference on Distributed Computing Systems, 2008, pp. 206–213.

[15] P. Indyk, H.Q. Ngo, A. Rudra, Efficiently decodable non-adaptive group testing, Proceedings of the 21st Annual ACM-SIAM Symposium on Discrete Algorithms, 2010, pp. 1126–1142.

[16] M. Cheraghchi, A. Karbasi, S. Mohajer, V. Saligrama, Graph-constrained group testing, Proceedings of International Symposium on Information Theory (ISIT), 2010, pp. 1913–1917.

[17] M. Mézard, C. Toninelli, Group testing with random pools: optimal two-stage algorithms, IEEE Transactions on Information Theory 57 (2011) 1736–1745.

[18] Y.-W. Hong, A. Scaglione, Group testing for sensor networks: the value of asking the right question, 38th Asilomar Conference on Signals, Systems and Computers, 2004, pp. 1297–1301.

[19] P. Braca, S. Marano, V. Matta, P. Willett, Asymptotic optimality of running consensus in testing statistical hypotheses, IEEE Transactions on Signal Processing 58 (2010) 814–825.

[20] R. Gallager, Low-density Parity-check Codes, Monograph, MIT Press, 1963.

[21] M. Cheraghchi, A. Hormati, A. Karbasi, M. Vetterli, Group testing with probabilistic tests: theory, design and application, IEEE Transactions on Information Theory 57 (2011) 7057–7067.

[22] T. Tošić, N. Thomos, P. Frossard, Distributed Sensor Failure Detection in Sensor Networks, Technical Report ⟨http://arxiv.org/abs/1109.5636⟩, 2011.