

A resilient consensus algorithm with inputs for the distributed monitoring of cyber-physical systems[☆]

Sabato Manfredi^{a,*}, David Angeli^{b,c}, Ciro Tortora^d

^a Department of Electrical Engineering and Information Technology of University of Naples Federico II, Naples, Italy

^b Department of Electrical and Electronic Engineering, Imperial College, London, UK

^c Department of Information Engineering, University of Florence, Florence, Italy

^d Terna Rete Italia S.p.A., Italy

ARTICLE INFO

Keywords:

Cyber-physical systems

Wireless sensor networks

Resilient distributed consensus algorithms

ABSTRACT

Recent advancements in multi-agent systems (MASs) have led to the development of numerous algorithms for achieving specific objectives, such as consensus. However, security remains a major challenge in MAS consensus, particularly addressing the adversarial behavior of malicious agents. This paper explores the extension of Mean-Subsequence-Reduced (MSR) algorithm-type mechanisms for resilient dynamic consensus in the presence of input reference signals. We provide necessary and sufficient conditions for resilient dynamic consensus without relying on the presence of trusted agents. Additionally, we experimentally validate the proposed algorithm and related conditions over a small cyber-physical system used for temperature monitoring. Furthermore, we propose and experimentally validate a fault-detection and recovery algorithm to achieve a resilient dynamic average consensus of regular agents.

1. Introduction

In recent years, there have been remarkable advancements in multi-agent systems (MASs). Numerous algorithms have been developed to achieve specific objectives using networks of intelligent agents, such as consensus (see, e.g., [Kia et al. \(2019\)](#) and references therein). One major challenge in MAS consensus is security, particularly addressing the adversarial behavior of malicious agents ([Ishii, Wang, & Feng, 2022](#); [Pirani, Mitra, & Sundaram, 2023](#); [Šubelj & Bajec, 2011](#)). These systems find applications in resilient rendezvous of robots to prevent them from being led to unsafe areas and in distributed computing to counteract the Byzantine adversary, which acts arbitrarily and may send malicious messages to its neighbors ([Ishii et al., 2022](#); [Park & Hutchinson, 2017](#)).

The problem of approximate Byzantine consensus with at most F adversaries was explored in [Dolev, Lynch, Pinter, Stark, and Weihl \(1986\)](#) using complete networks. The proposed resilient algorithm involved trimming the smallest and largest values and is referred to as Mean-Subsequence-Reduced (MSR) algorithms. This was later generalized to the Weighted-version (W-MSR) by Leblanc in [Šubelj and Bajec](#)

(2011), where the interplay between network topological robustness properties and consensus resilience was presented. Subsequent research has addressed resilient consensus problems in MASs from first-order to high-order dynamics or coupled vector-based systems ([Fu, Qin, Zheng, Chen, & Kang, 2021](#); [Sundaram & Gharesifard, 2019](#); [Usevitch & Panagou, 2020](#); [Wang & Ishii, 2020](#)).

In [Abbas, Shabbir, Li, and Koutsoukos \(2022\)](#), resilient consensus in high-dimensional systems was investigated. The primary idea involves normal agents updating their state iteratively by computing a 'safe point' within the convex hull of the initial states of normal agents. These methods are based on Tverberg partition or centerpoint in discrete geometry. Extensions to stochastic interaction scenarios are presented in [Shang \(2022\)](#), where single-input discrete-time controllable MASs are considered.

Besides pioneering works on discrete-time resilient consensus, continuous-time versions were considered in [LeBlanc and Koutsoukos \(2018\)](#), [Zhai, Liu, Guan, and Wen \(2021\)](#). Most approaches in the literature consider resilient static consensus algorithms, with only a few

[☆] This work was supported by project "Resilient Optimization in Distributed Cyber-Physical Control Problems subject to Adversarial Behaviour", funded in part by Italian Ministry of University and Research (MUR) within the Progetti di Rilevante Interesse Nazionale (PRIN) 2022 Program under Project 20225MESZB, and in part by European Union - Next Generation EU, Mission 4 Component 1, CUP E53D23000470006.

* Corresponding author.

E-mail addresses: sabato.manfredi@unina.it (S. Manfredi), d.angeli@imperial.ac.uk (D. Angeli).

¹ Part of the work was developed while Prof. S. Manfredi was a Visiting Academic with Control and Power Group at Electrical and Electronic Engineering Department, Imperial College, London.

extending to the presence of inputs, namely, to the resilient dynamic consensus framework.

Specifically, in [Iqbal, Qu, and Gusrialdi \(2022\)](#), a resilient dynamic consensus protocol was proposed to withstand false-data injection attacks modeled as disturbances. In contrast, [Bhattacharyya and Kalaimani \(2023\)](#) tackled resilient dynamic consensus in the face of Byzantine adversaries, representing a different form of adversarial attack compared to false-data injection. In [Bhattacharyya and Kalaimani \(2023\)](#), a novel resilient algorithm was introduced to provide state update laws for trusted and regular/normal agents to ensure the latter track the average of the reference signals of trusted agents. It was demonstrated that when trusted agents form a connected dominating set and the first-order differences of local reference signals are bounded, the proposed algorithm guarantees resilient dynamic average consensus of regular agents to the average value of reference signals of trusted agents. The presence of trusted agents removes limitations on the number of adversarial agents that may be present in the network. However, assuming the presence of trusted agents, as well as expecting regular agents to know these trusted agents beforehand, can be limiting in real-world applications. Trusted agents themselves are susceptible to attacks or faults. If attackers compromise or manipulate these trusted agents, they can undermine the entire system's resilience. Consequently, having a small number of trusted agents creates single or few points of failure. If any of these trusted agents fail or are compromised, it could disrupt the entire system's functionality.

Furthermore, relying on trusted agents can limit the scalability of the system. Managing and maintaining a sufficient number of trusted agents may become impractical or cost-prohibitive as the system grows.

Moreover, the approach where regular agents converge to a consensus value based on the reference signals of trusted agents might not be suitable for distributed monitoring systems. It may be more desirable for regular sensing agents to converge to the average value of their own measured reference signals, especially in scenarios where some agents are disrupted or compromised.

In this context, we explore the extension of MSR algorithm-type mechanisms for resilient dynamic average consensus in the presence of input reference signals, with the following improvements:

1. necessary and sufficient conditions for resilient dynamic consensus in the presence of inputs are formulated;
2. the proposed approach does not rely on the presence of trusted agents;
3. we experimentally validate the proposed fault/malicious detection algorithm as well as the related necessary and sufficient conditions over a small cyber-physical system (CPS) used for temperature monitoring. The CPS is a sensor network composed of off-the-shelf wireless communication nodes equipped with digital temperature sensors. This architecture has been widely implemented, mainly in industrial settings (e.g., [Manfredi \(2010\)](#));
4. a recovery algorithm is employed to complement the fault detection algorithm, ensuring the dynamic average consensus among the regular nodes. This guarantees accurate estimates of the average value of their input signals. The approach extends beyond relying solely on trusted agents, which only ensures convergence of regular agents to the average value of trusted agents' reference signals rather than their own. We experimentally validate the final fault-detection-recovery algorithm.

Notice that the inclusion of input signals in a system where faults and malicious nodes may exist indeed makes the problem more challenging compared to scenarios without input signals. Firstly, the algorithm's functionality is significantly extended to estimate the average values of input signals at nodes, which is more practically useful than estimating the average based solely on initial conditions. The algorithm allows for dynamic tracking of these average values, which

is particularly relevant for applications such as distributed cyber-physical monitoring systems. Secondly, in environments where faults and untrustworthy nodes are present, distinguishing between malicious behavior and legitimate input variations becomes more complex. Each agent's behavior is influenced by its input signal, which can naturally vary over time. This variation complicates the task of differentiating between legitimate deviations due to input signal changes and malicious attempts to disrupt the system. Malicious nodes might exploit these dynamics by manipulating their inputs to either blend in with normal operations or amplify their disruptive impact, making detection and mitigation more challenging. In contrast, systems without input signals typically exhibit more predictable dynamics, simplifying the identification of anomalies that may indicate malicious activity. From an analytical perspective, the presence of input signals means that the algorithm's dynamic model is no longer autonomous, making stability and convergence analysis more complex. The system's dynamics now involve both interactions among agents and external input signals, requiring the stability analysis to consider a wide range of input conditions, including worst-case scenarios where inputs might exacerbate the effects of malicious nodes. This scenario leads to non-uniform convergence across the network, as agents may converge to different values depending on their inputs, beyond their initial conditions. Thus, analyzing convergence in such systems necessitates accounting for the heterogeneity in input signals, further complicating the analysis.

The rest of the paper is organized as follows. In Section 2, the background on the robustness properties dealing with adversarial agents are introduced. Section 3 presents the proposed MSR algorithm with input reference signals along with the related theoretical necessary and sufficient conditions for its convergence. In Section 4, the experimental set-up and the experimental validation of the proposed algorithm, along with related theoretical results, are illustrated. Finally, conclusions and future works are given in Section 5.

2. Notation and graph robustness properties

Consider a dynamic network represented by directed graphs $G = G(V, E)$, where $V = \{1, \dots, n\}$ stands for the collection of nodes, and $E \subset V \times V$ symbolizes the set of directed edges. a_{ij} represents an element of the adjacency matrix of the graph A , where $a_{ij} = 1$ if there is a link from node j to node i , and $a_{ij} = 0$ otherwise. The value of a_{ij} is determined by the topology of the underlying physical graph of the network. This graph, which can be established a priori, dictates which nodes can physically communicate with each other based on factors such as physical distance, power transmission, and the communication protocol in use. The node collection is divided into two subsets: one is comprised of regular or normal nodes denoted as $\mathcal{N}$, and the other $V \setminus \mathcal{N}$ consists of adversary nodes, which remains unknown to the normal nodes in advance. Each directed edge $(i, j) \in E$ represents the flow of information, indicating that node i can potentially be influenced by (or receive information from) node j . The set of incoming neighbors, also known as neighbors, for node i is defined as $\mathcal{N}_i^i = \{j \in V : (i, j) \in E\}$, the in-degree of node i is denoted as $d_i = |\mathcal{N}_i^i|$ and is defined $d_M = \max_i d_i$. Similarly, the set of outgoing neighbors for node i is defined as $\mathcal{N}_i^o = \{j \in V : (j, i) \in E\}$. $\mathbb{N}_0$ denotes the set of natural numbers including 0.

In what follows, we will briefly recall the definitions that play a role in our analysis, originally introduced in ([LeBlanc, Zhang, Koutsoukos, & Sundaram, 2013](#)).

Definition 1 (F-Total Set). A set $S \subset V$ is F -total if it contains at most F nodes in the network, i.e., $|S| \leq F$, $F \in \mathbb{N}_0$.

Definition 2 (r-Reachable Set). Given a digraph G and a nonempty subset S of nodes of G , we say S is an r -reachable set if $\exists i \in S$ such that $|\mathcal{N}_i^i \setminus S| \geq r$, where $r \in \mathbb{N}_0$.

Definition 3. A set of adversary nodes is F -total bounded if it is an F -total set. We refer to these threat as the F -total model of deception attacks.

In essence, a set S is considered ' r -reachable' if it contains a node with no fewer than r neighbors outside of S . The parameter r quantifies the degree of redundant information flow from external nodes to an internal node within S . Put simply, ' r -reachability' characterizes the idea that within the set, at least one node is influenced by a sufficiently significant number of nodes from outside the set. This property specifically applies to the defined set S . Alternatively, another way to address this is by introducing a different form of information redundancy, which involves specifying a minimum number of nodes that experience sufficient influence from outside their respective sets. To capture this concept, we define the following.

Definition 4 ((r, s) -Reachable Set). Given a digraph G and a nonempty subset S of nodes of G , we say S is an (r, s) -reachable set if there are at least s elements $i \in S$ each of which has at least r neighbors outside of S , where $r, s \in \mathbb{N}_0$; i.e., given $X_S^r = \{i \in S : |\mathcal{N}_i^r \setminus S| \geq r\}$, then $|X_S^r| \geq s$.

Hence, (r, s) -reachability strictly generalizes r -reachability, and better quantifies the number of nodes with redundant information flow from outside of their set. This additional specificity is useful for defining a more general notion of robustness.

Definition 5 ((r, s) -Robustness). A nonempty, nontrivial digraph $G = G(V, E)$ on n nodes ($n \geq 2$) is (r, s) -robust, with $r \in \mathbb{N}_0$, $s \in [1, n]$ if for every pair of nonempty, disjoint subsets S_1, S_2 of V , at least one of the following holds:

1. (i) $|X_{S_1}^r| = |S_1|$
2. (ii) $|X_{S_2}^r| = |S_2|$
3. (iii) $|X_{S_1}^r| + |X_{S_2}^r| \geq s$

3. Resilient consensus algorithm with inputs

We consider, for each normal agent $i \in \mathcal{N}$, an extended version of the well-known mean subsequence reduced (MSR) algorithm that incorporates the input reference signal $r_i(k)$:

$$x_i(k+1) = x_i(k) + \varepsilon \sum_{j \in R_i(k)} a_{ij}(x_j(k) - x_i(k)) + \Delta r_i(k) \quad (1)$$

with $\Delta r_i(k) = r_i(k) - r_i(k-1)$, the step size $\varepsilon \in (0, \frac{1}{d_M})$ and $R_i(k)$ is a set built removing nodes from $\mathcal{N}_i^r$ according to the MSR rules. Namely, if less than F of the state values $x_j(k)$ of agents j in $\mathcal{N}_i^r$ are larger (resp. smaller) than $x_i(k)$ all such agents are removed from $\mathcal{N}_i^r$, else the agents with the F largest (resp. smallest) values are removed. Therefore, $R_i(k)$ is the subset of neighbors of agent i whose state values are received and retained by the node i according to the previous rules. This yields to the corresponding Algorithm 1 which will be referred to hereafter as the Dynamic MSR (D-MSR) algorithm.

If we need to track the average value of reference signals r_i , then the following distributed initialization at nodes is required: $x_i(0) = r_i(-1)$, $i \in \mathcal{N}$.

Assumption 1. The reference signals $r_i(k)$ are asymptotically convergent such that $|\Delta r_i(k)| \leq \Delta r$, $\forall k \in \mathbb{N}_0$ and some positive real Δr , and $\sum_{k=0}^{\infty} |\Delta r_i(k)| < \infty$, $\forall i$.

In the context of the considered scenarios with input signals, it is necessary for these inputs to satisfy Assumption 1, specifically that their variations are bounded and summable. Indeed, bounded variation ensures that the rate of change in input signals is limited, preventing abrupt or extreme changes that could destabilize the system or cause significant deviations in agent states. Summable variation implies that

Algorithm 1: D-MSR - Resilient consensus-based distributed MSR algorithm with inputs executed at each normal node i

Let parameters: $F, 0 < \varepsilon < \frac{1}{d_M}$
Initialize the estimation $x_i(0) = r_i(-1)$;
for $k = 0, 1, \dots$ **do**
 Acquire and store $r_i(k)$ from local sensor;
 Transmit $x_i(k)$ to its out-neighbors;
 Receive $x_j(k)$ from all of its in-neighbors;
 Order all the $x_j(k)$ received from smallest to largest;
 if Less than F of the received values are larger than $x_i(k)$ **then**
 Remove all the received values larger than $x_i(k)$;
 else
 Remove the F largest received values;
 end
 if Less than F of the received values are smaller than $x_i(k)$ **then**
 Remove all the received values smaller than $x_i(k)$;
 else
 Remove the F smallest received values;
 end
 Denote by $R_i(k)$ the subset of neighbors of agent i whose state values are received and retained by it;
 Calculate the quantity $\Delta r_i(k) = r_i(k) - r_i(k-1)$;
 Update the estimation as

$$x_i(k+1) = x_i(k) + \varepsilon \sum_{j \in R_i(k)} a_{ij}(x_j(k) - x_i(k)) + \Delta r_i(k)$$

end

the cumulative change in the input signals over time is finite, which is crucial for ensuring that the system can still reach a consensus or stable state, even as input signals vary. With summable variations, the impact of these changes diminishes over time, allowing the system to eventually settle into a stable state.

Definition 6 (Resilient Consensus with Inputs). Extending the original definition reported in the paper (LeBlanc et al., 2013) to the case of consensus protocols with inputs, we define the resilient consensus as characterized by the state of normal/regular agents fulfilling the following conditions:

- 1 **Agreement:** the states of normal agents converge to the same value $\bar{x}$ as k approaches infinity, i.e. $\lim_{k \rightarrow \infty} x_i(k) = \bar{x}$ for all $i \in \mathcal{N}$;
- 2 **Validity:** the states of normal agents reside for all $k \geq 0$ within the safety set: $[x_m(0) - \Delta r, x_M(0) + \Delta r]$

In the following we state the main result assessing the resilient consensus with inputs achieved by the algorithm (1).

We recall the following Corollary of Lemma 3 in Polyak (1987) that will be useful in the proof of the main result of Theorem 1.

Lemma 1. Let two non negative sequences of real numbers $\{u_k\}$ and $\{\beta_k\}$ such that $u_{k+1} \leq \alpha u_k + \beta_k$ for all $k \in \mathbb{N}_0$, if $\alpha \in [0, 1)$ and $\lim_{k \rightarrow \infty} \beta_k = 0$, then $\lim_{k \rightarrow \infty} u_k = 0$

In what follows, we will consider an undirected network graph $G = (V, E)$.

Theorem 1. Let an undirected network denoted by $G = (V, E)$ under F -total deception attacks such that each normal/regular node executes Algorithm 1 with parameter F and $\lim_{k \rightarrow \infty} \Delta r_i(k) = 0$ for all $i \in \mathcal{N}$. Then, the normal/regular nodes reach consensus for all choices of initial values and F -total sets of malicious nodes iff the graph is $(F+1, F+1)$ -robust.

Proof. We first prove the sufficiency part of the statement.

Sufficiency

Let $x_M(k) = \max_{i \in \mathcal{N}} x_i(k)$, $x_m(k) = \min_{i \in \mathcal{N}} x_i(k)$ and $D(k) = x_M(k) - x_m(k)$, $h \in \mathbb{N}_0$, $\delta \in \mathbb{R}$, we define the following set:

$$X_M(k, h, \delta) = \{i \in V : x_i(k+h) > x_M(k) - \delta\} \quad (2)$$

$$X_m(k, h, \delta) = \{i \in V : x_i(k+h) < x_m(k) + \delta\}. \quad (3)$$

Moreover, we let $\Delta r_M(k) = \max_i |\Delta r_i(k)|$ and the succession of real numbers $\Delta r_k = \sup_{\tau \geq k} \{\Delta r_M(\tau)\}$, $k \in \mathbb{N}_0$, such that $\Delta r_k \geq \Delta r_{k+1} \geq \Delta r_{k+2} \dots$

For any $k, h \in \mathbb{N}_0$ we denote $\delta(k, h) = \eta^h D(k)/2 - h\Delta r_k$ with

$$\eta = \min\{1 - \varepsilon \max_{i \in V} \sum_{j \in V} a_{ij}, \varepsilon \min_{i, j \in V: a_{ij} > 0} a_{ij}\},$$

Assuming the nontrivial case of positive $\delta(k, 0)$ for $h = 0$ we see that $X_M(k, h, \delta(k, 0))$ and $X_m(k, 0, \delta(k, 0))$ are non-empty, as they contain at least a regular node of value $x_M(k)$ and $x_m(k)$, respectively. Moreover, $\delta(k, 0) = D(0)/2$ and therefore $X_M(k, h, \delta(k, 0)) \cap X_m(k, 0, \delta(k, 0)) = \emptyset$. Since by assumption, the network is $(F+1, F+1)$ -robust with no more than F malicious nodes (i.e. F -total model), there is at least a normal node in $X_M(k, h, \delta(k, 0)) \cup X_m(k, 0, \delta(k, 0))$ that has at least $F+1$ neighbors outside of its respective set. Without loss of generality, let $X_M(k, 0, \delta(k, 0))$ be $F+1$ -reachable and a regular node $i \in X_M(k, 0, \delta(k, 0))$ be the node with at least $F+1$ neighbors outside $X_M(k, 0, \delta(k, 0))$. These neighbors have values at most equal to $x_M(k) - \delta(k, 0)$ so that at least one of them, say $\bar{j}$ will be used in its update by node i , since node i removes at most F values lower than its own value. Therefore from Algorithm 1 node i is affected by at least one neighbor $\bar{j} \in \mathcal{N}_i^i \setminus X_M(k, 0, \delta(k, 0))$. Therefore, we may infer:

$$\begin{aligned} x_i(k+1) &= [1 - \varepsilon \sum_{j \in R_i(k)} a_{ij}] x_i(k) + \varepsilon \sum_{j \in R_i(k) \setminus \{\bar{j}\}} a_{ij} x_j(k) \\ &\quad + \varepsilon a_{i\bar{j}} x_{\bar{j}}(k) + \Delta r_i(k) \leq [1 - \varepsilon \sum_{j \in R_i(k)} a_{ij}] x_M(k) \\ &\quad + \varepsilon \sum_{j \in R_i(k) \setminus \{\bar{j}\}} a_{ij} x_M(k) + \varepsilon a_{i\bar{j}} x_{\bar{j}}(k) + \Delta r_i(k) \\ &\leq (1 - \varepsilon a_{i\bar{j}}) x_M(k) + \varepsilon a_{i\bar{j}} (x_M(k) - \delta(k, 0)) \\ &\quad + \Delta r_i(k) = x_M(k) - \varepsilon a_{i\bar{j}} \delta(k, 0) + \Delta r_k \\ &\leq x_M(k) - \eta \delta(k, 0) + \Delta r_k = x_M(k) - \delta(k, 1). \end{aligned} \quad (4)$$

As a consequence $x_i(k+1) \in V \setminus X_M(k, 1, \delta(k, 1))$ and there exists at least a node $i \in X_M(k, 0, \delta(k, 0))$ falling in $V \setminus X_M(k, 1, \delta(k, 1))$ at step $k+1$. For nodes $j \in \mathcal{N} \setminus X_M(k, 0, \delta(k, 0))$ the following holds instead:

$$\begin{aligned} x_j(k+1) &= [1 - \varepsilon \sum_{i \in R_j(k)} a_{ji}] x_j(k) + \varepsilon \sum_{i \in R_j(k)} a_{ji} x_i(k) + \Delta r_j(k) \\ &\leq [1 - \varepsilon \sum_{i \in R_j(k)} a_{ji}] (x_M(k) - \delta(k, 0)) + \varepsilon \sum_{i \in R_j(k)} a_{ji} x_M(k) \\ &\quad + \Delta r_j(k) = x_M(k) - [1 - \varepsilon \sum_{i \in R_j(k)} a_{ji}] \delta(k, 0) + \Delta r_j(k) \\ &\leq x_M(k) - \eta \delta(k, 0) + \Delta r_k \leq x_M(k) - \delta(k, 1) \end{aligned}$$

Therefore normal nodes in the set $V \setminus X_M(k, 0, \delta(k, 0))$ at time k , belong to $V \setminus X_M(k, 1, \delta(k, 1))$ at time $k+1$. Hence by the above considerations we may infer that if $X_M(k, 0, \delta(k, 0))$ is $(F+1)$ reachable, it results:

$$|\mathcal{N} \cap X_M(k, 0, \delta(k, 0))| > |\mathcal{N} \cap X_M(k, 1, \delta(k, 1))|$$

Dually, if $X_m(k, 0, \delta(k, 0))$ is $(F+1)$ -reachable we may infer for a normal node $i \in X_m(k, 0, \delta(k, 0))$ the following inequality:

$$\begin{aligned} x_i(k+1) &\geq (1 - \eta) x_m(k) + \eta (x_m(k) + \delta(k, 0)) + \Delta r_i(k) \\ &\geq x_m(k) + \eta \delta(k, 0) - \Delta r_k = x_m(k) + \delta(k, 1) \end{aligned} \quad (5)$$

Similar inequality holds for the other normal nodes in $V \setminus X_m(k, h, \delta(k, 0))$, finally yielding:

$$|\mathcal{N} \cap X_m(k, 0, \delta(k, 0))| > |\mathcal{N} \cap X_m(k, 1, \delta(k, 1))|$$

Notice that, by definition, the succession $\delta(k, h)$ is monotonically decreasing w.r.t. $h \in \mathbb{N}_0$, i.e., $\delta(k, h+1) < \delta(k, h)$ for any $k, h \in \mathbb{N}_0$. Hence, $x_M(k) - \delta(k, h) > x_m(k) + \delta(k, h)$ holds for all $h > 0$ (since $x_M(k) - \delta(k, 0) \geq x_m(k) + \delta(k, 0)$ and, as a consequence the sets $\mathcal{N} \cap X_M(k, h, \delta(k, h))$ and $\mathcal{N} \cap X_m(k, h, \delta(k, h))$ are disjoint for all $h > 0$. Therefore the sets $\mathcal{N} \cap X_M(k, 1, \delta(k, 1))$ and $\mathcal{N} \cap X_m(k, 1, \delta(k, 1))$ are disjoint and, in addition, if both are not empty, from Definition 5, at least one of them is a $(F+1, F+1)$ -reachable set.

Assumed $\mathcal{N} \cap X_M(k, 1, \delta(k, 1)) \neq \emptyset$, being the graph $(F+1, F+1)$ robust, there exists at least a normal node i having $F+1$ neighbors outside $X_M(k, 1, \delta(k, 1))$. Let $\bar{j}$ a such neighbor, we may derive for $h = 2$:

$$\begin{aligned} x_i(k+2) &\leq [1 - \varepsilon \sum_{j \in R_i(k)} a_{ij}] x_M(k+1) + \varepsilon \sum_{j \in R_i(k) \setminus \{\bar{j}\}} a_{ij} x_M(k+1) \\ &\quad + \varepsilon a_{i\bar{j}} x_{\bar{j}} + \Delta r_i(k+1) \leq (1 - \varepsilon a_{i\bar{j}}) x_M(k+1) \\ &\quad + \varepsilon a_{i\bar{j}} (x_M(k) - \delta(k, 1)) + \Delta r_i(k+1) \\ &\leq (1 - \eta) x_M(k+1) + \eta x_M(k) - \eta \delta(k, 1) + \Delta r_M(k+1) \\ &\leq (1 - \eta) (x_M(k) + \Delta r_M(k)) + \eta x_M(k) - \eta^2 \delta(k, 0) \\ &\quad + \eta \Delta r_M(k) + \Delta r_M(k+1) = x_M(k) + \Delta r_M(k) \\ &\quad - \eta^2 \delta(k, 0) + \Delta r_M(k+1) \leq x_M(k) - \eta^2 \delta(k, 0) + 2\Delta r_k \\ &= x_M(k) - \delta(k, 2) \end{aligned}$$

where the latter inequality is derived considering that:

$$\Delta r_k \geq \max\{\Delta r_M(k), \Delta r_M(k+1)\}.$$

Hence, there exists at least a normal node $i \in X_M(k, 1, \delta(k, 1))$ falling in $V \setminus X_M(k, 2, \delta(k, 2))$ at step $k+2$.

For the other normal nodes $j \in V \setminus X_M(k, 1, \delta(k, 1))$ a similar inequality holds such that they stay in $V \setminus X_M(k, 1, \delta(k, 1))$ at step $k+1$, while they stay in $V \setminus X_M(k, 2, \delta(k, 2))$ at step $k+2$.

If $\mathcal{N} \cap X_m(k, 1, \delta(k, 1)) \neq \emptyset$ and there exists at least a normal agent i that is $F+1$ reachable, applying similar arguments it results:

$$\begin{aligned} x_i(k+2) &\geq x_m(k) - \Delta r_M(k) + \eta^2 \delta(k, 0) - \Delta r_M(k+1) \\ &\geq x_m(k) + \eta^2 \delta(k, 0) - 2\Delta r_k = x_m(k) + \delta(k, 2) \end{aligned}$$

By iteration at each subsequent time step $k+h$, $h \geq 3$, if $\mathcal{N} \cap X_m(k, 1, \delta(k, 1)) \neq \emptyset$ and $\mathcal{N} \cap X_M(k, 1, \delta(k, 1)) \neq \emptyset$, then one of them will decrease its cardinality. In addition, if either of these sets becomes empty, it will be empty in the subsequent h iterations as any normal node outside of these sets will either be constrained by an upper bound of $x_M(k) - \delta(k, h)$ or a lower bound of $x_m(k) + \delta(k, h)$. Hence after $|\mathcal{N}|$ time steps, at least one of the sets $\mathcal{N} \cap X_M(k, h, \delta(k, h))$ or $\mathcal{N} \cap X_m(k, h, \delta(k, h))$ will become empty.

Assuming the former set will become empty, it results:

$$\begin{aligned} x_M(k+|\mathcal{N}|) &\leq x_M(k) - \delta(k, |\mathcal{N}|) = x_M(k) - \eta^{|\mathcal{N}|} \delta(k, 0) \\ &\quad + |\mathcal{N}| \Delta r_k = x_M(k) - \eta^{|\mathcal{N}|} \frac{D(k)}{2} + |\mathcal{N}| \Delta r_k \end{aligned}$$

In addition, being $x_m(k+|\mathcal{N}|) \geq x_m(k) - |\mathcal{N}| \Delta r_k$ we may infer:

$$D(k+|\mathcal{N}|) \leq \left(1 - \frac{\eta^{|\mathcal{N}|}}{2}\right) D(k) + 2|\mathcal{N}| \Delta r_k \quad (6)$$

The same expression arises if the set $\mathcal{N} \cap X_m(k, h, \delta(k, h))$ becomes empty.

Notice that over each time window $[h|\mathcal{N}|, (h+1)|\mathcal{N}|]$, if $D(h|\mathcal{N}|) > 0$ (i.e. consensus is not achieved at time $k = h|\mathcal{N}|$) at least the sets $\mathcal{N} \cap X_M(h|\mathcal{N}|, 0, \delta(h|\mathcal{N}|, 0))$ and $\mathcal{N} \cap X_m(h|\mathcal{N}|, 0, \delta(h|\mathcal{N}|, 0))$ are not empty with at least one of them $(2F+1)$ -reachable. Hence, iterating the above procedure over each time window $[h|\mathcal{N}|, (h+1)|\mathcal{N}|]$ with $\delta(h|\mathcal{N}|, 0) = D(h|\mathcal{N}|)/2$, from (6) and generality of index k , we may infer:

$$D((k+1)|\mathcal{N}|) \leq \left(1 - \frac{\eta^{|\mathcal{N}|}}{2}\right) D(k|\mathcal{N}|) + 2|\mathcal{N}| \Delta r_{k|\mathcal{N}|}$$

Being $1 - \frac{\eta^{|\mathcal{N}|}}{2} \in (0, 1)$ and by the assumption $\limsup_{k \rightarrow \infty} \Delta r_M(k) = 0$ yielding $\lim_{k \rightarrow \infty} \Delta r_{k|\mathcal{N}|} = 0$, and finally by Lemma 1 it follows

the convergence to zero of the considered subsequence $\{D(k|\mathcal{N})\}_k$, namely, $\lim_{k \rightarrow \infty} D(k|\mathcal{N}) = 0$.

From the generality of the considered starting point of the subsequence, the above convergence result holds as well for all the $|\mathcal{N}| - 1$ subsequences $\{D(k|\mathcal{N}|h)\}_k$, obtained for $h = 1, \dots, |\mathcal{N}| - 1$. This is enough to assess that: $\lim_{k \rightarrow \infty} D(k) = 0$. Then the states of normal agents converge (exponentially) to a consensus equilibrium. Being $\Delta r_k \leq \Delta r$, by (4) and (5) the following holds for any $k \in \mathbb{N}_0$ and normal agents $i \in \mathcal{N}$:

$$x_i(k+1) \in [x_m(k) - \Delta r, x_M(k) + \Delta r] \quad (7)$$

Therefore, throughout all time intervals the states of normal agents reside within the safety set $[x_m(0) - \Delta r, x_M(0) + \Delta r]$. Notice that if we assume bounded reference signals with $|r_i(k)| \leq \hat{r}$ for all k , the safety set boils down to $[-x_m(0) - 2\hat{r}, x_M(0) + 2\hat{r}]$, being $|\Delta r| \leq 2\hat{r}$.

Necessity

Now we prove the necessity part of the statement. To this aim, we assume the network does not fulfill $(F+1, F+1)$ -robustness and prove that nodes cannot always make an agreement.

If the network is not $(F+1, F+1)$ -robust, we can identify two disjoint sets, $S_1 \subset V$ and $S_2 \subset V$, non $(F+1)$ -reachable and such that none of conditions 1. - 3. of Definition 5 hold. Therefore the negation of conditions 1. and 2. yields existence of some node in S_1 and in S_2 with at most F neighbors outside its respective set. In addition, negation of condition 3. implies at most F nodes in $S_1 \cup S_2$ having at least $F+1$ neighbors outside their respective set. We assume such F nodes in $S_1 \cup S_2$ being adversarial, and forming a network F -total set according to Definition 1. We assume specific constant reference signals for normal nodes within S_1 such that $r_i(k) = r_1$ for $i \in S_1, k \geq -1$. Similarly, normal nodes within S_2 read constant reference signals such that $r_i(k) = r_2 > r_1, i \in S_2, k \geq -1$.

In addition, assume nodes $i \in V \setminus (S_1 \cup S_2)$ are associated to constant reference signals: $r_i(k) \in [r_1, r_2], k \geq -1$. From the initialization of the algorithm, we choose $x_i(0) = r_1, x_i(0) = r_2, x_i(0) \in [r_1, r_2]$ for normal nodes in S_1, S_2 and $V \setminus (S_1 \cup S_2)$, respectively. Furthermore, we assume malicious adversarial nodes in S_1 and S_2 being initialized and assuming a constant value equal to r_1 and r_2 , respectively for all $k \geq -1$.

The normal agents in S_2 are connected at most to F nodes outside S_2 . Therefore the algorithm (1) filters F values coming from either S_1 or $V \setminus (S_1 \cup S_2)$. In addition as the malicious nodes in S_2 keep a constant value equal to r_2 , the normal nodes in S_2 never deviate from their initial value r_2 . Similar arguments can be exploited for S_1 , yielding $x_i(k) = r_1$ for $i \in S_1, k \geq 1$. Moreover, all nodes in $V \setminus (S_1 \cup S_2)$ starting in $[r_1, r_2]$ and applying algorithm (1) are a convex combinations of elements in $[r_1, r_2]$, and hence belong to this interval for all k . In summary, the normal nodes in S_1 and S_2 will never deviate from their distinct initial values, and thus consensus will not be reached for the considered scenario of a graph not fulfilling $(F+1, F+1)$ -robustness. $\square$

4. Experimental validation

4.1. Experimental set-up

We validate the proposed algorithm over a cyber-physical system composed of wireless sensor nodes equipped with a temperature sensor. The architecture can be used for distributed temperature monitoring. The considered architecture is represented in Fig. 1. Specifically, each sensor node consists of a sensing unit, e.g. digital temperature sensor DHT-11, a communication unit, XBee RF module, and a computation unit based on Arduino microcontroller. The communication takes place at 2.4 GHz according to ZigBee protocol. In addition, one or more nodes are equipped with an Arduino Wifi so as to upload data (e.g. the calculated estimates x_i and reference signals r_i defined by the measured temperatures) of the nodes on the IoT services platform Thingspeak.

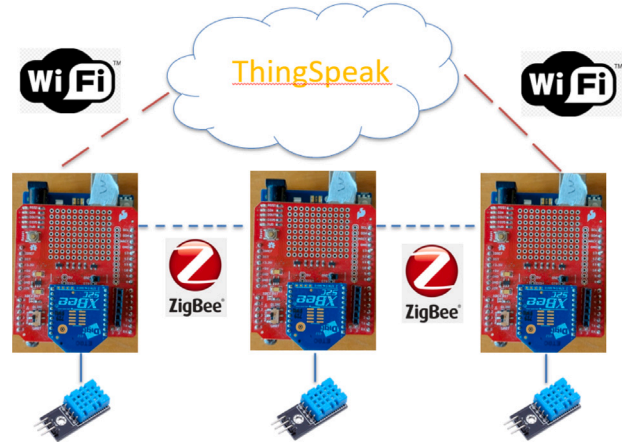


Fig. 1. Architecture of the considered cyber-physical system for temperature monitoring.

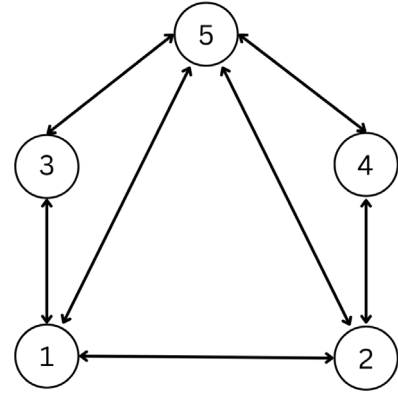


Fig. 2. 5-node topology (2,2)-robust.

4.2. Detection and recovery algorithm experimental validation

In this section, we are going to use the hardware presented above to set up a cyber-physical system and experimentally validate the theoretical results and Algorithm 1. From now on, all practical experiments have been carried out considering the topology shown in Fig. 2, that is (2,2)-robust. The tuning of ϵ was carefully done to balance the need for speed with the need for stability and accuracy while fulfilling the condition $\epsilon < \frac{1}{d_M}$, ultimately setting $\epsilon = 0.1$.

To compare the performance of the standard dynamic consensus algorithm (Kia et al., 2019), D-MSR Algorithm 1, and Algorithm 1 with the Recovery phase (later defined as D-MSR²), we evaluate the following index:

$$e_M[\%] = \max_{i \in \mathcal{N}} \frac{|x_i(\infty) - \bar{r}|}{|\bar{r}|} \cdot 100 \quad (8)$$

where $x_i(\infty)$ represents the final value of the state variable x_i for any normal agent after reaching its steady state. $e_M[\%]$ represents the maximum relative error (in percentage) between any agent's final state and the desired consensus value. The aim is to quantify how far the worst-performing normal agent in the network deviates from the average value of reference signals of normal nodes $\bar{r}$. In an ideal scenario, $e_M[\%]$ would be 0%, meaning all agents perfectly reach the consensus value $\bar{r}$. A higher $e_M[\%]$ indicates a larger deviation and poorer performance of the considered algorithm. Additionally, for D-MSR Algorithm 1, we evaluate the Safety Interval (i.e., the interval within which the estimates x_i of normal agents are guaranteed to reside), derived from

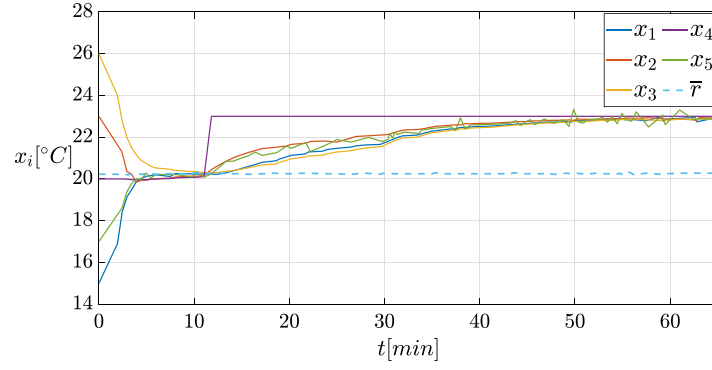


Fig. 3. Average temperature estimations x_i (continous lines) provided by the Standard Dynamic Consensus algorithm (Kia et al., 2019) in the presence of STEP malicious/faulty behavior of agent 4 (dashed-dot line).

Table 1

Performance evaluation and comparison of the standard dynamic consensus algorithm and D-MSR Algorithm 1 under STEP, RAMP, and PERIODICAL-type fault/malicious agent behavior.

Experiment scenario (Figure)	Standard consensus (Kia et al., 2019) e_M [%]	Safety interval from bound (7)	D-MSR Algorithm 1 e_M [%]
STEP (Fig. 4)	13.9	[14.7–26.3]	2.3
RAMP (Fig. 6)	∞	[16.5–26.5]	4.2
PERIODICAL (Fig. 7)	9	[16.7–26.3]	6

Table 2

Performance evaluation and comparison of the standard dynamic consensus, D-MSR Algorithm 1 and D-MSR² algorithm under STEP, RAMP, and PERIODICAL-type fault/malicious agent behavior.

Experiment scenario (Figure)	Standard consensus (Kia et al., 2019) e_M [%]	Safety interval from bound (7)	D-MSR Algorithm 1 e_M [%]	D-MSR ² Algorithm e_M [%]
STEP (Fig. 9)	9.3	[16.5–26.5]	3.2	0.6
RAMP (Fig. 10)	∞	[16.0–27.0]	1.4	0.2
PERIODICAL (Fig. 11)	23.2	[16.5–26.5]	3	0.6

bound (7), in accordance with the validity condition 2 of the Resilience Definition 6. The results are reported on Tables 1 and 2.

Finally, to validate the effectiveness of the proposed algorithms, we consider scenarios involving three representative types of attacks or disruptions, where the malicious or faulty node at a given step injects erroneous data following step, ramp, and periodical signals, denoted as STEP, RAMP, and PERIODICAL in Tables 1 and 2, respectively.

Firstly we implemented the standard dynamic average consensus protocol (Kia et al., 2019) over a network in Fig. 2, demonstrating why it fails when a node in the network no longer adhere to the protocol due to faults or malicious attacks. In Fig. 3, we observe that when node 4 deviates from the protocol by injecting values according to a step signal, it can maliciously or erroneously steer the rest of the network toward its own value, as node 4 effectively assumes the role of a leader. In the experiment under discussion, node 4 deviates from the protocol by maintaining a constant value while continuing to influence the final equilibrium through the communication network. Indeed, strong and symmetric (or at least balanced) network connectivity is essential for the dynamic consensus protocol to accurately estimate the average of the inputs, as is the case with the topology shown in Fig. 2. Hence, given the connected nature of the network, node 4 effectively drives all nodes toward its value, acting as a leader. While this behavior can be predicted if the anomalous dynamics of node 4 (or any other node) are known in advance in terms of timing and behavior, the challenge arises when such deviations occur unexpectedly, as in the case of malicious attacks or faults. Therefore, due to the inherent structure of the consensus protocol and its reliance on connectivity assumptions, the negative effects of node 4 (or similar nodes) are, unfortunately, unavoidable. In this case where all normal agents follow the malicious/faulty node 4, we get a value of $e_M = 13.7\%$ for the normal agent. Based on the above discussion, when node 4 injects ramp or periodical signals, the normal nodes eventually follow these patterns, with $e_M = \infty$ occurring in the case of the ramp signal.

Subsequently, we validate the proposed resilient dynamic MSR consensus protocol (1) by implementing the related D-MSR Algorithm 1 over the CPS. We consider three different types of malicious attacks or faults.

The first scenario, illustrated in Fig. 4, involves node 4 having a step variation. We observe, according to the theoretical results, that the normal nodes $i = 1, 2, 3, 5$ achieve consensus.

The second scenario involves node 4 behaving as a ramp either occurring after resilient consensus is reached (Fig. 5), or while the normal nodes are approaching to the consensus equilibrium (Fig. 6). In both cases the algorithm makes the normal nodes resilient, allowing them to achieve consensus.

The final type of attack or fault is the trickiest to detect or reject, represented by a periodical waveform with maximum (resp. minimum) larger (smaller) to the state values of normal nodes.

In Fig. 7 is shown that even if the attacker infiltrates the range of normal nodes during time, the malicious agent can no longer disrupt them.

Table 1 presents a comparison between the standard dynamic consensus algorithm and the D-MSR algorithm in the scenarios shown in Figs. 4, 6 and 7. The results confirm that D-MSR algorithm achieves lower error, maintaining it finite in the case of ramp signal injection. Therefore D-MSR Algorithm 1 is efficient as a fault/malicious detection algorithm, enabling consensus among normal nodes. However, it is not guaranteed that the final consensus value is the desired estimation of the average value of temperature reference signals of normal nodes (e.g., $\bar{r}$, dashed line in the Figures). Specifically, the validity condition from Definition 6 only guarantees that the normal agents' estimates x_i remain within the safety interval derived from bound (7), without providing any guarantees regarding their steady state consensus value. This bound can be conservative, as indicated by the numerical evaluation of the safety interval in Table 1.

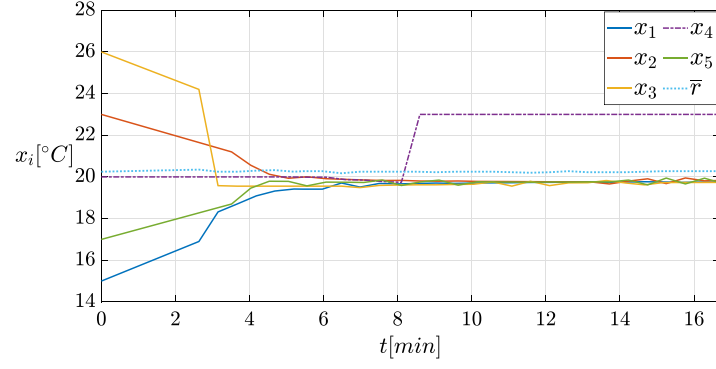


Fig. 4. Average temperature estimations x_i (continuous lines) provided by D-MSR Algorithm 1 in the presence of STEP malicious/faulty behavior of agent 4 (dashed-dot line).

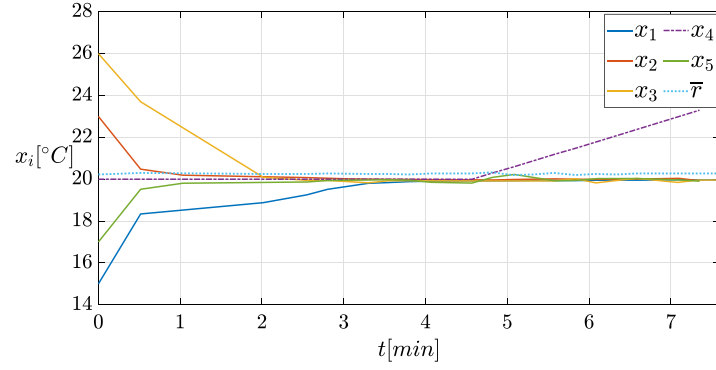


Fig. 5. Average temperature estimations x_i (continuous lines) provided by D-MSR Algorithm 1 in the presence of RAMP malicious/faulty behavior of agent 4 (dashed-dot line).

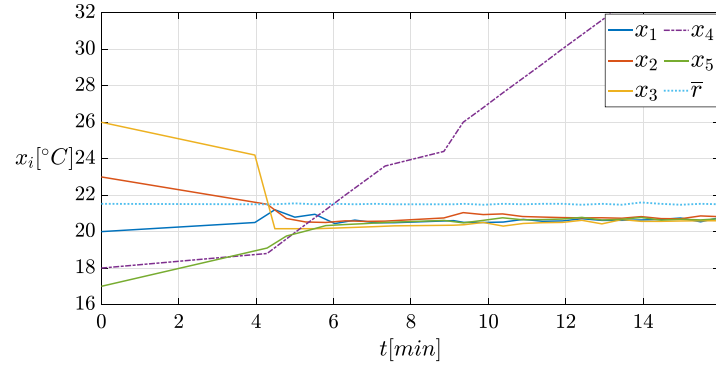


Fig. 6. Average temperature estimations x_i (continuous lines) provided by D-MSR Algorithm 1 in the presence of early RAMP malicious/faulty behavior of agent 4 (dashed-dot line).

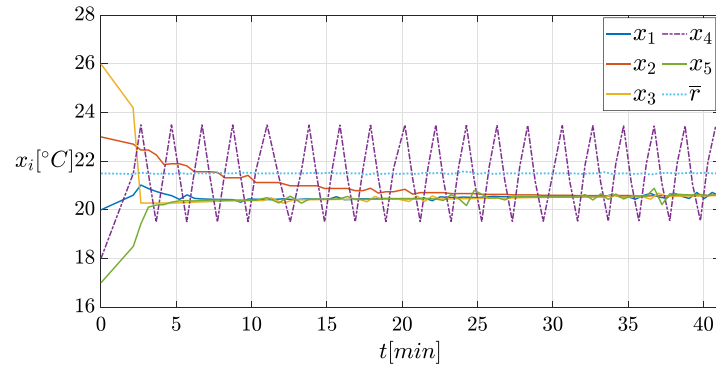


Fig. 7. Average temperature estimations x_i (continuous lines) provided by D-MSR Algorithm 1 in the presence of PERIODICAL malicious/faulty behavior of agent 4 (dashed-dot line).

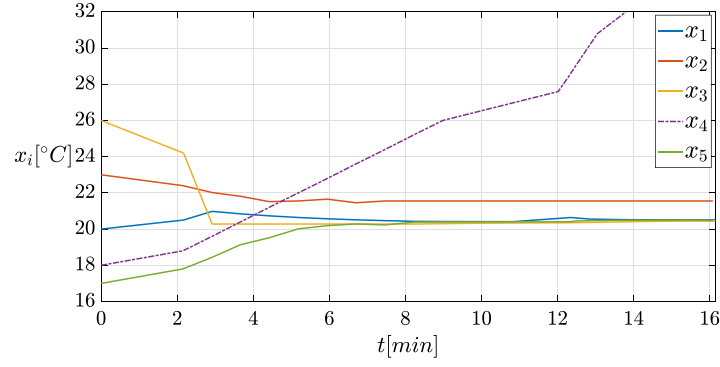


Fig. 8. Average temperature estimations x_i (continous lines) provided by D-MSR Algorithm 1 in the presence of malicious/faulty behavior of agent 4 (dashed-dot line) for a network with lower robustness than (2,2) one.

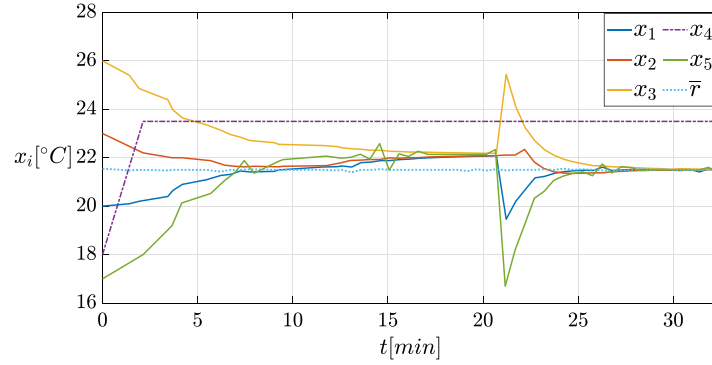


Fig. 9. Average temperature estimations x_i (continous lines) provided by D-MSR² algorithm (Detection and Recovery phase) in the presence of STEP malicious/faulty behavior of agent 4 (dashed-dot line).

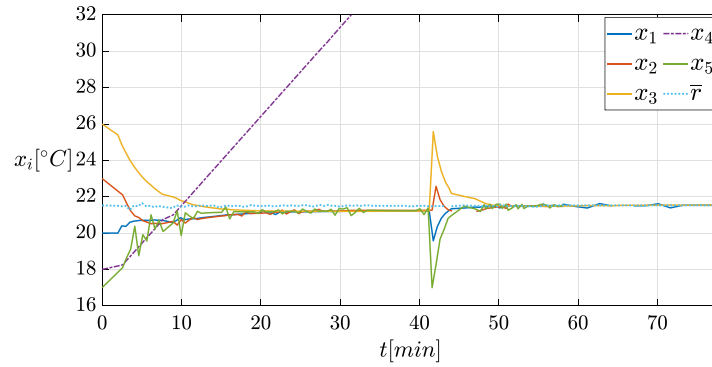


Fig. 10. Average temperature estimations x_i (continous lines) provided by D-MSR² algorithm (Detection and Recovery phase) in the presence of RAMP malicious/faulty behavior of agent 4 (dashed-dot line).

Finally, we validate the necessary condition of [Theorem 1](#) by modifying the topology shown in [Fig. 2](#) such that the hypotheses of [Theorem 1](#) are no longer fulfilled, and the network is no more (2,2) robust. [Fig. 8](#) experimentally confirms that in a such situation, with $F = 1$, the normal nodes are no longer able to reach a resilient consensus, with, specifically, node 2 unable to achieve agreement with the other normal agents.

So far, we have shown that when the conditions of [Theorem 1](#) are satisfied, normal/regular nodes can indeed achieve resilient consensus and implement fault/malicious detection operation, effectively filtering out the attackers. However, this algorithm, when employed alone, disrupts the conservativeness property, essential to guarantee the correct tracking of the average value of the regular agents' reference signals. To address this, we propose firstly utilizing D-MSR Algorithm 1 to detect potential faults or malicious agents within the network. Subsequently,

we apply an implementation of the dynamic asynchronous consensus algorithm, which is distributively executed at the regular nodes after detecting the faults or malicious agents, to attain the correct estimation of the average value of their reference signals.

In the following, we will outline the key features of the latter algorithm that constitutes the recovery phase, referring to the original works ([Demetri, Olfati-Saber, & Richard, 2005](#); [Kriegledler, Oung, & D'Andrea, 2013](#)) for further details. The algorithm offers an alternative implementation of the standard dynamic consensus algorithm, capable of handling time-varying topologies while preserving the conservativeness property. This is crucial to ensure the convergence of the estimations to the exact average value of the reference signals r_i .

This is achieved by introducing auxiliary variables δ_{ij} for each neighbor j of agent i , which store the cumulative disagreement between the two agents. This results in the following algorithm implementation:

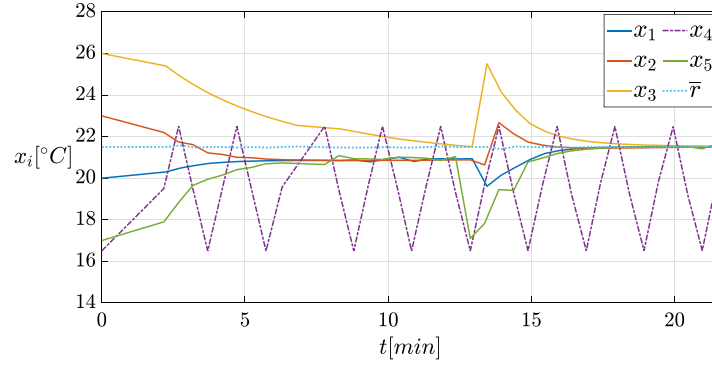


Fig. 11. Average temperature estimations x_i (continuous lines) provided by D-MSR² algorithm (Detection and Recovery phase) in the presence of PERIODICAL malicious/faulty behavior of agent 4 (dashed-dot line).

$$\begin{aligned} \delta_{ij}(k+1) &= \delta_{ij}(k) - a_{ij}(x_i(k) - x_j(k)) \\ x_i(k+1) &= \epsilon \sum_{j=1}^N \delta_{ij}(k+1) + r_i(k) \end{aligned} \quad (9)$$

Specifically, when an agent leaves the network, its neighbors simply set the corresponding δ_{ij} to zero, effectively neutralizing its contribution. Consequently, the states will converge to the appropriate average of the r_i values in the modified network. The key idea is to leverage the algorithm in (9) to dynamically and asynchronously construct the network of normal agents, which have reached consensus as a result of the detection phase of D-MSR Algorithm 1, to accurately estimate the average value $\bar{r}$ of their reference signals r_i , $i \in \mathcal{N}$. Specifically, each normal agent i sets $\delta_{ij} = 0$ for any neighbor j with which it is not in agreement and has been identified as malicious or faulty, as determined by D-MSR Algorithm 1 (i.e., the detection phase).

In doing so, the fault detection/isolation phase of D-MSR Algorithm 1 is complemented by a recovery phase (9) to correctly estimate the average value of the reference signals of regular nodes, thus enhancing the robustness and applicability of the protocol to distributed monitoring cyber-physical systems. The resulting D-MSR Algorithm 1, along with the Recovery phase (9), is referred to as D-MSR².

We validate the overall fault detection-recovery algorithm D-MSR² in similar scenarios as previously presented. Results are shown in Figs. 9, 10, and 11, confirming that eventually the regular agents well estimate the average value of their reference input temperature signals (e.g., $\bar{r}$, dashed line in the Figures). Note that the sharp changes observed in Figures 9–11 at times 21, 41, and 13, respectively, are due to the activation of the recovery phase of D-MSR² algorithm.

For comparison, Table 2 reports the performance evaluation of the standard dynamic consensus algorithm, D-MSR, and D-MSR². The results confirm that D-MSR² outperforms both the standard dynamic consensus algorithm and D-MSR, achieving smaller final error. The final error is small but not zero, due to the unavoidable imprecision of the sensors and network communication in the experimental setup.

In the context of the evaluated consensus algorithms — namely, the standard dynamic consensus algorithm, D-MSR, and D-MSR² — *resilience* refers to the system's ability to ensure that normal agents can continue to compute accurate or bounded estimates, even under adversarial conditions (i.e. the algorithm effectively resists the influence of adversarial conditions). Tables 1 and 2 highlight the resilient consensus achieved by the D-MSR and D-MSR² algorithms, which exhibit a final estimation error that is finite and smaller than that of the standard consensus algorithm. D-MSR guarantees that normal nodes maintain finite, bounded estimates in the presence of adversarial agents, even though the final consensus value may not reflect the average of the normal nodes' reference signals. D-MSR² further enhances this capability by reducing the final error, thereby improving the accuracy of the proposed resilient strategy.

Remark 1. In the absence of malicious nodes or faults, the convergence time of D-MSR² algorithm is indeed higher than that of the standard protocol. This is primarily due to the successive application of the fault detection and recovery phases. Additionally, during the detection phase, the filtering operation may slow down the convergence time since it effectively removes links. However, this effect is negligible for small values of F . Moreover, this impact can be mitigated by designing a more connected or robust network topology. Since the number of steps for the consensus, detection, and recovery phases does not exceed the network diameter, the additional time due to the recovery phase can be estimated to be no greater than the network diameter.

5. Conclusions and future work

This paper has presented an extension of Mean-Subsequence-Reduced (MSR) algorithm-type mechanisms for resilient dynamic consensus in the presence of input reference signals in cyber-physical systems. By providing necessary and sufficient conditions for resilient dynamic consensus without relying on the presence of trusted agents, we have addressed a significant limitation in existing approaches. Furthermore, the experimental validation of the proposed algorithm and related conditions over a cyber-physical system used for temperature monitoring demonstrates its practical effectiveness.

Moreover, we have proposed and experimentally validated a fault-detection and recovery algorithm to achieve resilient dynamic average consensus among the regular agents, enhancing the applicability to distributed CPS monitoring systems. These contributions advance the state-of-the-art in resilient consensus algorithms for MASs, providing a more comprehensive framework for achieving dynamic average consensus in the presence of faults and adversarial behavior. Ongoing research is focused on extending the validation of the proposed approaches to industrial settings and to applications beyond temperature monitoring.

CRediT authorship contribution statement

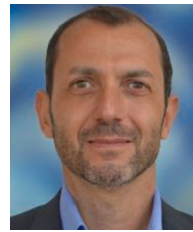
Sabato Manfredi: Writing – review & editing, Writing – original draft, Validation, Supervision, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization. **David Angeli:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization. **Ciro Tortora:** Validation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

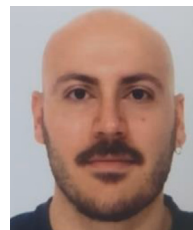
- Abbas, W., Shabbir, M., Li, J., & Koutsoukos, X. (2022). Resilient distributed vector consensus using centerpoint. *Automatica*, 136, Article 110046.
- Bhattacharyya, S., & Kalaimani, R. K. (2023). Resilient dynamic average consensus based on trusted agents. *arXiv:2303.08171*.
- Demetri, P. S., Olfati-Saber, R., & Richard, M. M. (2005). Dynamic consensus on mobile networks. In *IFAC world congress*.
- Dolev, D., Lynch, N. A., Pinter, S. S., Stark, E. W., & Weihl, W. E. (1986). Reaching approximate agreement in the presence of faults. *Journal of the ACM*, 33(3), 499–516.
- Fu, W., Qin, J., Zheng, W. X., Chen, Y., & Kang, Y. (2021). Resilient cooperative source seeking of double-integrator multi-robot systems under deception attacks. *IEEE Transactions on Industrial Electronics*, 68(5), 4218–4227.
- Iqbal, M., Qu, Z., & Gusrialdi, A. (2022). Resilient dynamic average-consensus of multiagent systems. *IEEE Control Systems Letters*, 6, 3487–3492.
- Ishii, H., Wang, Y., & Feng, S. (2022). An overview on multi-agent consensus under adversarial attacks. *Annual Reviews in Control*, 53, 252–272.
- Kia, S. S., Van Scoy, B., Cortes, J., Freeman, R. A., Lynch, K. M., & Martinez, S. (2019). Tutorial on dynamic average consensus: The problem, its applications, and the algorithms. *IEEE Control Systems Magazine*, 39(3), 40–72.
- Kriegleder, M., Oung, R., & D'Andrea, R. (2013). Asynchronous implementation of a distributed average consensus algorithm. In *2013 IEEE/RSJ international conference on intelligent robots and systems* (pp. 1836–1841).
- LeBlanc, H. J., & Koutsoukos, X. (2018). Resilient first-order consensus and weakly stable, higher order synchronization of continuous-time networked multiagent systems. *IEEE Transactions on Control of Network Systems*, 5(3), 1219–1231.
- LeBlanc, H. J., Zhang, H., Koutsoukos, X., & Sundaram, S. (2013). Resilient asymptotic consensus in robust networks. *IEEE Journal on Selected Areas in Communications*, 31(4), 766–781.
- Manfredi, S. (2010). A reliable cooperative and distributed management for wireless industrial monitoring and control. *International Journal of Robust and Nonlinear Control*, 20(2), 123–139. <http://dx.doi.org/10.1002/rnc.1453>.
- Park, H., & Hutchinson, S. A. (2017). Fault-tolerant rendezvous of multirobot systems. *IEEE Transactions on Robotics*, 33(3), 565–582.
- Pirani, M., Mitra, A., & Sundaram, S. (2023). Graph-theoretic approaches for analyzing the resilience of distributed control systems: A tutorial and survey. *Automatica*, 157, Article 111264.
- Polyak, B. (1987). *Introduction to Optimization*. New York: Optimization Software.
- Shang, Y. (2022). Median-based resilient consensus over time-varying random networks. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 69(3), 1203–1207.
- Šubelj, L., & Bajec, M. (2011). Unfolding communities in large complex networks: Combining defensive and offensive label propagation for core extraction. *Phys. Rev. E*, 83, Article 036103.
- Sundaram, S., & Gharesifard, B. (2019). Distributed optimization under adversarial nodes. *IEEE Transactions on Automatic Control*, 64(3), 1063–1076.
- Usevitch, J., & Panagou, D. (2020). Resilient leader-follower consensus to arbitrary reference values in time-varying graphs. *IEEE Transactions on Automatic Control*, 65(4), 1755–1762.
- Wang, Y., & Ishii, H. (2020). Resilient consensus through event-based communication. *IEEE Transactions on Control of Network Systems*, 7(1), 471–482.
- Zhai, Y., Liu, Z.-W., Guan, Z.-H., & Wen, G. (2021). Resilient consensus of multi-agent systems with switching topologies: A trusted-region-based sliding-window weighted approach. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 68(7), 2448–2452.



Sabato Manfredi is currently an Associate Professor of automatic control with the Department of Electrical Engineering and Information Technology, University of Naples Federico II, Italy. He was a Contingent Worker with the Control and Power Group at the Department of Electrical and Electronic Engineering, Imperial College London, from November 2012 to December 2022, and has continued as an Academic Visitor since then. He was also a Visiting Professor at the School of Mathematical Sciences, Queen Mary University of London, from 2016 to 2017. He has authored/co-authored more than 100 scientific publications including the monograph “Multilayer Control of Networked Cyber-Physical Systems”, *Advances in Industrial Control Series*, Springer, 2017. He collaborates with many international universities and companies, holds European patent and is involved in a range of academic, industrial and consulting projects. His research interests include distributed estimation, control and optimization of/over cyber-physical systems and complex networks, decentralized machine learning techniques for intelligent networked multi-agent systems.



David Angeli received the B.S. degree in computer science engineering and the Ph.D. degree in control theory from University of Florence, Italy, in 1996 and 2000, respectively. Since 2000, he was with the Department of Systems and Computer Science, University of Florence, as an Assistant Professor, and an Associate Professor, in 2005. In 2007, he was a Visiting Professor with I.N.R.I.A de Rocquencourt, Paris, France, and he joined as a Senior Lecturer with the Department of Electrical and Electronic Engineering, Imperial College London, London, U.K., in 2008, where he is currently a Professor in Nonlinear Network Dynamics and the Director of Postgraduate Teaching. He has authored more than 100 journal papers in the areas of stability of nonlinear systems, control of constrained systems (MPC), chemical reaction networks theory, and smart grids. Prof Angeli is a Fellow of the IEEE since 2015, and of the Institution of Engineering and Technology, since 2018. He has served as an Associate Editor for IEEE TRANSACTIONS IN AUTOMATIC CONTROL AND AUTOMATICA.



Ciro Tortora received a master's degree in automation engineering from the University of Naples Federico II and a second-level master's degree in digitalization of the electricity system for the energy transition from the University of Salerno in 2022 and 2024, respectively. He currently works for the Italian transmission system operator Terna in the Scada systems development and digitalization of electrical substations unit.