

Realtime ℓ_1 -fault-and-state estimation for multi-agent systems

Kazumune Hashimoto, Michelle Chong and Dimos V. Dimarogonas

Abstract—In this paper, we propose a state and fault estimation scheme for multi-agent systems. The proposed estimator is based on an ℓ_1 -norm optimization problem, which is inspired by sparse signal recovery in the field of compressive sampling. Moreover, we provide a necessary and sufficient condition such that state and fault signals are correctly estimated. The result presents a fundamental limitation of the algorithm, which shows how many faulty nodes are allowed to ensure a correct estimation. An illustrative example of a vehicle platoon is given to validate the effectiveness of the proposed approach.

I. INTRODUCTION

The control of multi-agent systems arises in many areas, such as vehicle platooning, formation control of autonomous vehicles and satellites, cooperative control of robots and so on [1]. In recent years, the design of a framework for detecting anomaly or faults has attracted much attention, see e.g. [2]. This is motivated by the fact that multi-agent systems typically consist of many dynamical systems interacting with each other. Hence, only a few anomalous behaviors can result in a significant impact on the overall system's performance. For example, in vehicle platooning, anomalous behaviors such as the sudden braking or acceleration of a leading vehicle will directly affect the follower vehicles, which may cause accidents or congestions. Thus, the detection and estimation of faults in a multi-agent system at an early stage is crucial to maintain safe and reliable real-time operations.

So far, various formulations and approaches have been proposed to detect and estimate faults in multi-agent systems, see, e.g., [3]–[9]. In this paper, we are particularly interested in developing a fault estimation strategy for multi-agent systems, where each node is able to measure the relative state information with respect to its neighbors. Some approaches have already been proposed in this problem set-up, see, e.g., [7]–[9], where the references include both centralized and distributed (decentralized) schemes. For example, in [9] a coordinate transformation is introduced to extract the observable subspace, and a sliding mode observer is designed to estimate faulty signals. In [8], a fault detection scheme is developed by introducing a set-valued observer, which is applied to two subsystems that are decomposed by left-coprime factorization. In addition, an event-triggered communication protocol based on unknown input observers was proposed in [7], where each node monitors the neighboring nodes for detecting faults. One remaining theoretical analysis of the above approaches is to guarantee the resilience of the estimator when multiple nodes are subject to faults. For example, the results presented in [7]–[9] neither provide a systematic procedure to estimate faults when multiple nodes are faulty, nor analyze how many faulty nodes can be tolerated to provide a correct estimation. Thus, the estimator may fail when multiple nodes are subject to

faults, and, even though the estimator is subject to failure, it is of great importance to analyze when this happens.

The contribution of this paper is to present a new fault estimation scheme for multi-agent systems based on relative state measurements. The objective is to provide a real-time estimator for both *states and faults*. Thus, this objective differs from some previous works in [7]–[9], where only fault signals are detected or estimated. As we will see later, this is achieved by considering the case when a node assigned as a leader is able to measure additional information. The proposed estimator is different from previous works in the literature, which utilizes an ℓ_1 -norm optimization problem inspired by sparse signal recovery in compressive sampling [10]. The optimization problem is placed in the Basis Pursuit [11], in which several numerical solvers can be employed to solve the problem efficiently. The estimator requires a centralized communication, in which each node assigned as a follower needs to transmit the measurement to the leader. Nevertheless, the estimator is a unified framework that can handle the case when multiple faults arise, which may be advantageous over the previous results in [7]–[9]. In addition, we provide a quantitative analysis of when the approach provides a correct estimation. The result provides a fundamental limitation of the algorithm, which illustrates how many faulty nodes can be tolerated such that both states and faults are correctly estimated.

The remainder of this paper is organized as follows. In Section II, we present some preliminaries on graph theory and an ℓ_1 -norm optimization problem. In Section III, the problem formulation and the estimator are given. In Section IV, theoretical analysis of the estimator is given. In Section V, an illustrative example shows the effectiveness of the proposed approach. Conclusions are given in Section VI.

Notation: Let \mathbb{R} , \mathbb{R}_+ , \mathbb{N} , \mathbb{N}_+ be the non-negative reals, positive reals, non-negative integers, and positive integers, respectively. For a given $x = [x_1, x_2, \dots, x_n] \in \mathbb{R}^n$, let $\|x\|_1$ be the ℓ_1 -norm of x , i.e., $\|x\|_1 = |x_1| + |x_2| + \dots + |x_n|$. For a given $M \in \mathbb{N}_+$, denote by $1_M \in \mathbb{R}^M$ a M -dimensional vector whose components are all 1. For a given set T , denote by $|T|$ the cardinality of T . For given $T \subseteq \{1, \dots, n\}$ and $x = [x_1, x_2, \dots, x_n] \in \mathbb{R}^n$, x is called T -sparse if all components indexed by the complement of T is 0, i.e., $x_i = 0, \forall i \in \{1, \dots, n\} \setminus T$. For a given $x \in \mathbb{R}^n$ and a set $T \subseteq \{1, \dots, n\}$, denote by $x_T \in \mathbb{R}^{|T|}$ a vector from x by extracting all components indexed by T . Given $A \in \mathbb{R}^{m \times n}$, denote by $\text{rank}(A)$ the rank of A . Moreover, denote by $\ker(A)$ the null-space of A , i.e., $\ker(A) = \{x \in \mathbb{R}^n : Ax = 0\}$.

II. PRELIMINARIES

A. Graph theory

Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ denote a directed graph, where $\mathcal{V} = \{1, 2, \dots, M\}$ is the set of nodes and $\mathcal{E} = \{e_1, e_2, \dots, e_N\} \subseteq \mathcal{V} \times \mathcal{V}$ is the set of edges. The graph is weakly connected if for every pair of nodes, when removing all orientations in the graph, there exists a path between them. For a given $i \in \mathcal{V}$, let $\mathcal{N}_i \subset \mathcal{V}$ be the set of neighboring nodes of i , i.e., $\mathcal{N}_i = \{j \in \mathcal{V} : e = \{i, j\} \in \mathcal{E}\}$. The incidence matrix $D = D(\mathcal{G})$ is the $\{0, \pm 1\}$ matrix, where $D^{(i,j)} = 1$ ($D^{(i,j)}$ denotes the (i, j) -component of D) if the node i is the head of the edge e_j , $D^{(i,j)} = -1$ if the node i is the tail of the edge e_j , and $D^{(i,j)} = 0$ otherwise. For a given $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with $|\mathcal{V}| = M$, it is shown that $\text{rank}(D) = M - 1$ if \mathcal{G} is weakly connected. The null-space of the incidence matrix is given by $\ker(D^T) = \gamma \mathbf{1}_M$, where $\gamma \in \mathbb{R}$ (see, e.g., Chapter 2 in [12]).

B. The Null-Space Property and ℓ_1 -norm optimization

In what follows, we provide the notion of the Null-Space Property (NSP) and its useful result for the ℓ_1 -norm optimization problem.

Definition 1 (The Null-Space Property). For a given $A \in \mathbb{R}^{m \times n}$ and $T \subseteq \{1, 2, \dots, n\}$, A is said to satisfy the Null-Space Property (NSP) for T (or T -NSP for short), if for every $v \in \ker(A) \setminus \{0\}$, it holds that

$$\|v_T\|_1 < \|v_{T^c}\|_1, \quad (1)$$

where $T^c = \{1, \dots, n\} \setminus T$. \square

The NSP is a key property to check whether the sparse signal can be reconstructed based on the ℓ_1 -norm optimization problem:

Theorem 1 (ℓ_1 -reconstruction theorem). For a given $A \in \mathbb{R}^{m \times n}$ and $T \subseteq \{1, \dots, n\}$, every T -sparse vector $x_0 \in \mathbb{R}^n$ is a unique solution to the following optimization problem:

$$\min_x \|x\|_1 \quad \text{s.t.} \quad Ax_0 = Ax, \quad (2)$$

if and only if A satisfies T -NSP. \square

The above theorem indicates that every T -sparse vector can be reconstructed by solving the ℓ_1 -norm optimization problem in (2), if and only if the matrix A satisfies T -NSP. The proof of Theorem 1 follows the same line as [13], [14] and is omitted in this paper.

III. PROBLEM FORMULATION

A. Sensing and communication architectures

Let us first provide the sensing and communication architectures considered in this paper. Consider a multi-agent system consisting of M nodes that are labeled by $\{1, 2, \dots, M\}$. The *sensing architecture* is modeled as a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{1, 2, \dots, M\}$ is the set of nodes and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is the set of edges. Here, each edge $\{i, j\} \in \mathcal{E}$ indicates the sensing capacity of node i with respect to node j . Specifically, if $\{i, j\} \in \mathcal{V}$, node i is able to measure relative state information

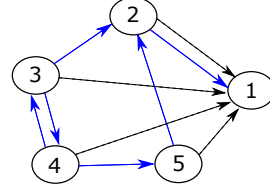


Fig. 1. Sensing and communication architectures considered in this paper. Blue arrows indicate the sensing model and black dotted arrows indicate the communication model.

with respect to j by using on-board sensors. We assume that the graph \mathcal{G} is weakly connected. The *communication architecture* is given as follows. Assume that there exists one leader and $M - 1$ followers; without loss of generality, we consider that the node labeled with 1 is the leader and the others are the followers. Then, each follower is capable of transmitting its measurements (obtained by the sensor) to the leader over the communication network. That is, the leader collects all measurements from the followers for each time, and estimates the states and fault signals of all nodes to monitor their behaviors. The sensing and communication architectures are illustrated in Fig. 1.

B. Dynamics and measurement model

The dynamics for node i is described by the following Linear Time-Invariant (LTI) system:

$$x_i(k) = A_i x_i(k-1) + B_i u_i(k-1) + f_i(k), \quad (3)$$

for $k \in \mathbb{N}$, where $x_i \in \mathbb{R}^n$ is the state, $u_i \in \mathbb{R}^m$ is the control input, and $f_i \in \mathbb{R}^n$ is the signal indicating the occurrence of faults, i.e., $f_i(k) \neq 0$ when node i is subject to fault at time step k , and $f_i(k) = 0$ otherwise. The faulty signals can be viewed as system faults on the dynamics in (3), or actuator faults caused by physical effects or cyber attacks over the communication network. Note that the state and input dimensions are n and m for all nodes. We assume that the control input u_i is known to node i for all $k \in \mathbb{N}$. Let $\mathcal{I}_k \subseteq \{1, \dots, M\} \cup \emptyset$ be the *unknown* set of faulty nodes at $k \in \mathbb{N}$, i.e.,

$$\mathcal{I}_k = \{i \in \{1, \dots, M\} : f_i(k) \neq 0\}. \quad (4)$$

Note that the set \mathcal{I}_k can be time-varying, i.e., the set of faulty nodes is possibly changing according to the time evolution.

The measurement outputs depend on whether the node is the leader or one of the followers. For the followers, each node is able to measure the *relative state information* from its neighbors by using on-board sensors, i.e., for all $i \in \mathcal{V} \setminus \{1\}$,

$$y_{ij}(k) = x_i(k) - x_j(k), \quad j \in \mathcal{N}_i \quad (5)$$

for $k \in \mathbb{N}$, where $y_{ij}(k)$, $j \in \mathcal{N}_i$ is the measurement output. The set of outputs collected at the node i is then expressed as

$$y_i(k) = C_i x(k), \quad (6)$$

where $y_i(k) \in \mathbb{R}^{n|\mathcal{N}_i|}$ is the vector collecting the measurement outputs for all $j \in \mathcal{N}_i$, C_i is the matrix defined through (5),

and $x(k) = [x_1(k)^\top \ x_2(k)^\top \ \dots \ x_M(k)^\top]^\top$.

Regarding the outputs for the leader, we consider the following two cases. First, it can measure not only the relative state information from its neighbors, but also its own state information, i.e.,

$$y_{1j}(k) = x_1(k) - x_j(k), \quad j \in \mathcal{N}_1, \quad (7)$$

$$y_{11}(k) = x_1(k), \quad (8)$$

where $y_{1j}(k)$, $j \in \mathcal{N}_1$ and $y_{11}(k)$ are the measurement outputs. If the leader is able to obtain both (7) and (8), we say that the leader is in *active mode*. Second, it can measure, similarly to the followers, only the relative state information from its neighbors, i.e.,

$$y_{1j}(k) = x_1(k) - x_j(k), \quad j \in \mathcal{N}_1, \quad (9)$$

and then we say that the leader is in *non-active mode*. In practice, the leader's mode indicated above changes over time depending on the physical environments. For example, if the leader moves in the outdoor environment that can utilize the GPS to locate the state, the measurement in (8) is available and thus the leader becomes active. On the other hand, if the leader enters an indoor environment in which it cannot utilize the GPS due to signal loss by the presence of walls in the building (e.g., inside the tunnel, underground, etc.), the measurement in (8) is *not* available and the leader is non-active. To indicate whether the leader is active or non-active, let $a_1(k) \in \{0, 1\}$ be given by

$$a_1(k) = \begin{cases} 0, & \text{if the leader is in nonactive mode} \\ 1, & \text{if the leader is in active mode.} \end{cases}$$

Regarding the leader's mode, we further assume the following.

Assumption 1. The leader is in the active mode at the initial time, i.e., $a_1(0) = 1$. Moreover, there is no fault at $k = 0$, i.e., $\mathcal{I}_0 = \emptyset$. \square

Note that the leader can be either active or non-active for all $k \in \mathbb{N} \setminus \{0\}$. As we will see in the analysis of Section IV, the assumption of being an active mode at $k = 0$ leads to a correct fault and state estimation for all times afterwards. As with (6), the set of measurement outputs for the leader can be expressed as

$$y_1(k) = C_{10}x(k), \quad \text{if } a_1(k) = 0, \quad (10)$$

$$y_1(k) = C_{11}x(k), \quad \text{if } a_1(k) = 1, \quad (11)$$

where C_{10} is the matrix obtained by collecting all measurements from (9), and C_{11} is the matrix obtained from (7) and (8). Based on (3) and (6) the overall dynamics of all nodes can be expressed as

$$x(k) = Ax(k-1) + Bu(k-1) + f(k), \quad (12)$$

for $k \in \mathbb{N}$, where

$$x(k) = [x_1(k)^\top \ x_2(k)^\top \ \dots \ x_M(k)^\top]^\top, \quad (13)$$

$$u(k) = [u_1(k)^\top \ u_2(k)^\top \ \dots \ u_M(k)^\top]^\top, \quad (14)$$

$$f(k) = [f_1(k)^\top \ f_2(k)^\top \ \dots \ f_M(k)^\top]^\top, \quad (15)$$

and A, B are the matrices with appropriate dimensions. The output is given by

$$y(k) = \begin{cases} C_0x(k), & \text{if } a_1(k) = 0 \\ C_1x(k), & \text{if } a_1(k) = 1, \end{cases} \quad (16)$$

where $y(k) = [y_1(k)^\top \ y_2(k)^\top \ \dots \ y_M(k)^\top]^\top$ and

$$C_0 = \begin{bmatrix} C_{10} \\ C_2 \\ \vdots \\ C_M \end{bmatrix}, \quad C_1 = \begin{bmatrix} C_{11} \\ C_2 \\ \vdots \\ C_M \end{bmatrix}. \quad (17)$$

C. State and fault estimator design

For each $k \in \mathbb{N}$, all followers obtain the measurement outputs according to (6) and transmit them to the leader over the communication network. Then, the leader estimates both the states of all nodes $x(k)$ and the fault signals $f(k)$, based on the knowledge of the dynamics and measurement outputs in (12) and (16). Let $\hat{x}(k)$, $\hat{f}(k)$, $k \in \mathbb{N}$ be the state and the fault signals that are estimated at k , respectively. In this paper, we propose the following estimator to obtain these estimates:

$$\begin{aligned} \hat{x}(k) &= \arg \min_x \|x - A\hat{x}(k-1) - Bu(k-1)\|_1 \\ \text{s.t. } y(k) &= Cx, \end{aligned} \quad (18)$$

where $C = C_0$ if $a_1(k) = 0$ and $C = C_1$ if $a_1(k) = 1$. Without loss of generality, we set $\hat{x}(k-1) = 0$ for $k = 0$. The estimation for $\hat{f}(k)$ is then computed as

$$\hat{f}(k) = \hat{x}(k) - A\hat{x}(k-1) - Bu(k-1). \quad (19)$$

In general, the term $A\hat{x}(k-1) + Bu(k-1)$ is called *apriori estimate* for k . Thus, the optimization problem in (18) aims at finding the state that is the closest to the apriori estimate, subject to the constraint on the measurement outputs. Letting $z = x - A\hat{x}(k-1) - Bu(k-1)$, the problem in (18) leads to

$$\min_x \|z\|_1 \quad \text{s.t. } y'(k) = Cz, \quad (20)$$

where $y'(k) = y(k) - C(A\hat{x}(k-1) + Bu(k-1))$, which is indeed a well-known Basis Pursuit (BP) formulation [11]. Hence, various numerical solvers can be applied to obtain the solution to (20), see, e.g., [15].

IV. ANALYSIS OF THE ESTIMATOR

In this section we provide a theoretical analysis of the proposed estimator. Suppose that Assumption 1 holds ($a_1(0) = 1$, $\mathcal{I}_0 = \emptyset$), and consider the unknown sets of faulty nodes $\mathcal{I}_k \subseteq \{1, \dots, M\} \cup \emptyset$, for $k \in \mathbb{N} \setminus \{0\}$, which means that $f_i(k) \neq 0, \forall i \in \mathcal{I}_k$ and $f_i(k) = 0$ otherwise. Let $\mathcal{F}_{\mathcal{I}_k} \subset \mathbb{R}^{nM}$ be given by

$$\mathcal{F}_{\mathcal{I}_k} = \{f(k) : f_i(k) \neq 0, \forall i \in \mathcal{I}_k \text{ and } f_i(k) = 0, \forall i \notin \mathcal{I}_k\}.$$

That is, $\mathcal{F}_{\mathcal{I}_k}$ is the domain of all $f(k)$ when the set of faulty nodes is given by \mathcal{I}_k . Then, we can say that the estimator is *correct* if for all $k \in \mathbb{N} \setminus \{0\}$ the optimization problem in (18)

yields the estimation that is identical to the actual one, i.e., for all $k \in \mathbb{N} \setminus \{0\}$, it follows that $\hat{x}(k) = x(k)$, $\hat{f}(k) = f(k)$, $\forall f(k) \in \mathcal{F}_{\mathcal{I}_k}$, $\forall u(k) \in \mathbb{R}^{mM}$. The following result presents a necessary and sufficient condition for this as the main result of this paper:

Theorem 2. Consider the estimator in (18), and suppose that Assumption 1 holds. Then, the following statements are equivalent:

- (a) $|\mathcal{I}_k| < M/2$, $\forall k \in \mathbb{N} \setminus \{0\}$.
- (b) $\hat{x}(k) = x(k)$, $\hat{f}(k) = f(k)$, $\forall f(k) \in \mathcal{F}_{\mathcal{I}_k}$, $\forall u(k-1) \in \mathbb{R}^{mM}$, $\forall a_1(k) \in \{1, 0\}$, $\forall k \in \mathbb{N} \setminus \{0\}$. \square

Theorem 2 indicates that state and fault signals are correctly estimated for all $k \in \mathbb{N} \setminus \{0\}$ regardless of the values of $f(k)$, control inputs $u(k-1)$ and the leader's mode $a_1(k)$, if and only if the number of faulty nodes is less than half the total number of nodes for all $k \in \mathbb{N} \setminus \{0\}$. In other words, it does not provide a correct estimation if the number of faulty nodes is more than or equal to $M/2$. Indeed, this fundamental limitation is related to the ones presented in [16]–[18], where they showed that the number of attacks on sensors that is tolerated to provide a correct estimation cannot exceed half the total number of sensors.

While Theorem 2 exhibits some similarities to the previous works in [16]–[18], our result is different in the following sense. The previous results consider a reconstructability of states under sensor attacks for only *observable* systems. In contrast, as described in Section III-B, we include here the case when the overall system becomes *un-observable* due to the constraint that only relative state information is available if the leader is non-active. As we will see below, this result can be shown by deriving an explicit condition to satisfy the null-space property of the measurement matrix C .

In order to provide the proof of Theorem 2, we resort to the following lemma:

Lemma 1. For a given $k \in \mathbb{N} \setminus \{0\}$, suppose that $\hat{x}(k-1) = x(k-1)$ holds. Then, the following statements are equivalent:

- (a) $|\mathcal{I}_k| < M/2$.
- (b) $\hat{x}(k) = x(k)$, $\hat{f}(k) = f(k)$, $\forall f(k) \in \mathcal{F}_{\mathcal{I}_k}$, $\forall u(k-1) \in \mathbb{R}^{mM}$, $\forall a_1(k) \in \{1, 0\}$. \square

Lemma 1 indicates that if there is no estimation error at $k-1$, a correct estimation is given at k if and only if $|\mathcal{I}_k| < M/2$. Let us provide the proof of Lemma 1. To this end, it is required to take several steps to modify the optimization problem in (18). Since $x(k-1) - \hat{x}(k-1) = 0$, we have $x(k) = A\hat{x}(k-1) + Bu(k-1) + f(k)$ and so $y(k) = C(A\hat{x}(k-1) + Bu(k-1) + f(k))$. Thus, the optimization problem becomes $\min \|x - A\hat{x}(k-1) - Bu(k-1)\|_1$, subject to $C(A\hat{x}(k-1) + Bu(k-1) + f(k) - x) = 0$. Letting $z = x - A\hat{x}(k-1) - Bu(k-1)$ be the new decision variable to obtain

$$\min_z \|z\|_1 \quad \text{s.t.} \quad Cf(k) = Cz, \quad (21)$$

where $C = C_0$ if $a_1(k) = 0$ and $C = C_1$ if $a_1(k) = 1$. Let $\hat{z}(k)$ be the solution to (21). Since $\hat{f}(k)$ is given by

(19) and $z = x - A\hat{x}(k-1) - Bu(k-1)$, we have $\hat{z}(k) = \hat{f}(k)$. Let $z = [z_1^T, z_2^T, \dots, z_M^T]^T$, where $z_i \in \mathbb{R}^n$, $i \in \{1, \dots, M\}$ represents the variable for node i . Then, we rearrange the vectors z , $f(k)$ and the matrix C in the following way. Let $T_i \subset \{1, \dots, nM\}$, $i \in \{1, 2, \dots, n\}$ be given by $T_i = \{i, i+n, i+2n, \dots, i+(M-1)n\}$, and rearrange z as $z_p = [z_{p_1}^T, z_{p_2}^T, \dots, z_{p_n}^T]^T$, where $z_{p_i} = z_{T_i}$, $i \in \{1, 2, \dots, n\}$. Roughly speaking, z_{p_i} represents a vector collecting the i -th component of all the nodes. Similarly, rearrange $f(k)$ as $f_p = [f_{p_1}^T, f_{p_2}^T, \dots, f_{p_n}^T]^T$, where $f_{p_i} = f_{T_i}$, $i \in \{1, 2, \dots, n\}$. Note that due to the rearrangement of $f(k)$, $f_i(k) \neq 0$ implies that $(f_p)_{\tilde{T}_i} \neq 0$, where $\tilde{T}_i = \{i, i+M, i+2M, \dots, i+(n-1)M\}$. Thus, letting $\tilde{\mathcal{F}}_{\mathcal{I}_k} \subset \mathbb{R}^{nM}$ be given by

$$\tilde{\mathcal{F}}_{\mathcal{I}_k} = \{f_p(k) \in \mathbb{R}^{nM} : (f_p(k))_{\tilde{T}_i} \neq 0, \forall i \in \mathcal{I}_k \text{ and } (f_p(k))_{\tilde{T}_i} = 0, \forall i \notin \mathcal{I}_k\}, \quad (22)$$

it follows that $f(k) \in \mathcal{F}_{\mathcal{I}_k} \iff f_p(k) \in \tilde{\mathcal{F}}_{\mathcal{I}_k}$. Note that $f_p(k) \in \tilde{\mathcal{F}}_{\mathcal{I}_k}$ implies that $f_p(k)$ is \tilde{T} -sparse, where $\tilde{T} = \cup_{i \in \mathcal{I}_k} \tilde{T}_i$. Moreover, by rearranging the columns of C according to the rearrangement for z as well as suitably rearranging the rows, one can always construct the matrix C_{p0} for the active mode ($a_1(k) = 1$) and C_{p1} for the non-active mode ($a_1(k) = 0$) as follows:

$$C_{p0} = \begin{bmatrix} D_0 & 0 & \cdots & 0 \\ 0 & D_0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & D_0 \end{bmatrix}, C_{p1} = \begin{bmatrix} D_1 & 0 & \cdots & 0 \\ 0 & D_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & D_1 \end{bmatrix},$$

where $D_0 = D^T$ and $D_1 = [e, D]^T$ with D being the incidence matrix of the graph \mathcal{G} and $e = [1, 0, 0, \dots, 0]^T \in \mathbb{R}^M$. Based on the above rearrangement, the optimization problem becomes

$$\min_{z_p} \|z_p\|_1 \quad \text{s.t.} \quad C_p f_p(k) = C_p z_p, \quad (23)$$

where $C_p = C_{p0}$ if $a_1(k) = 0$ and $C_p = C_{p1}$ if $a_1(k) = 1$. Let $\hat{f}_p(k)$ be the optimal solution of z_p to (23). Note that the problem is invariant under the above rearrangements. That is, letting $\hat{z}(k) = \hat{f}(k)$ and $\hat{f}_p(k) = [\hat{f}_{p_1}(k)^T, \hat{f}_{p_2}(k)^T, \dots, \hat{f}_{p_n}(k)^T]^T$ be the solutions to (21) and (23), respectively, it follows that $\hat{f}_{p_i}(k) = \hat{f}_{T_i}(k)$, $\forall i \in \{1, \dots, M\}$, i.e., the optimal solutions to (21) and (23) are equivalent under the above rearrangements. Thus, we obtain

$$f(k) = \hat{f}(k), \forall f(k) \in \mathcal{F}_{\mathcal{I}_k} \iff f_p(k) = \hat{f}_p(k), \forall f_p(k) \in \tilde{\mathcal{F}}_{\mathcal{I}_k}. \quad (24)$$

Based on the above notations and calculations, the proof of Lemma 1 is provided as follows.

(Proof of Lemma 1): (a) \Rightarrow (b) : Suppose that $|\mathcal{I}_k| < M/2$ holds. For the active mode $a_1(k) = 1$, we have $C_p = C_{p1}$ and since the graph is weakly connected, it is shown that $\text{rank}(C_p) = nM$ and $\text{ker}(C_p) = 0$. This means that z_p satisfying the constraint $C_p(f_p(k) - z_p) = 0$ is a *single point* and is given by $z_p = f_p(k)$, which shows that the optimization problem in (23) yields $\hat{f}_p(k) = f_p(k)$. This implies from (24)

that we have $\hat{f}(k) = f(k)$, $\forall f(k) \in \mathcal{F}_{\mathcal{I}_k}$, $\forall u(k-1) \in \mathbb{R}^{mM}$. Moreover, since $\hat{f}(k) = f(k)$, it follows that

$$\begin{aligned}\hat{x}(k) &= A\hat{x}(k-1) + Bu(k-1) + \hat{f}(k) \\ &= Ax(k-1) + Bu(k-1) + f(k) \\ &= x(k),\end{aligned}\quad (25)$$

which means that the state $x(k)$ is also correctly estimated. Thus, for the active mode case, state and fault signal are correctly estimated by solving (18). Let us now consider the non-active mode case $C_p = C_{p0}$. Let $\tilde{T}_i = \{i, i+M, i+2M, \dots, i+(n-1)M\}$, $i \in \{1, \dots, M\}$ and $\tilde{T} = \cup_{i \in \mathcal{I}_k} \tilde{T}_i$. Since $C_p = C_{p0}$ and \mathcal{G} is weakly connected, we obtain

$$\ker(C_p) = [\gamma_1 1_M^T, \gamma_2 1_M^T, \dots, \gamma_n 1_M^T]^T, \quad (26)$$

where $\gamma_i \in \mathbb{R}$, $i \in \{1, \dots, n\}$ and we have used the null-space of the incidence matrix D (see Section II-A). Thus, it follows that $\|v_{\tilde{T}_i}\|_1 = \sum_{j=1}^n |\gamma_j|$ for any $v \in \ker(C_p)$. Thus, we obtain

$$\|v_{\tilde{T}}\|_1 = \sum_{i \in \mathcal{I}_k} \|v_{\tilde{T}_i}\|_1 = |\mathcal{I}_k| \left(\sum_{j=1}^n |\gamma_j| \right) \quad (27)$$

$$\|v_{\tilde{T}^c}\|_1 = (M - |\mathcal{I}_k|) \left(\sum_{j=1}^n |\gamma_j| \right), \quad (28)$$

where $\tilde{T}^c = \{1, 2, \dots, nM\} \setminus \tilde{T}$. Since $|\mathcal{I}_k| < M/2$, it follows that $\|v_{\tilde{T}}\|_1 < \|v_{\tilde{T}^c}\|_1$, $\forall v \in \ker(C_p) \setminus \{0\}$, i.e., C_p satisfies \tilde{T} -NSP. Since $f_p(k) \in \tilde{\mathcal{F}}_{\mathcal{I}_k}$ implies that $f_p(k)$ is \tilde{T} -sparse, it follows from Theorem 1 that $f_p(k) = \hat{f}_p(k)$, $\forall f_p(k) \in \tilde{\mathcal{F}}_{\mathcal{I}_k}$. This implies that $\hat{f}(k) = f(k)$, $\forall f(k) \in \mathcal{F}_{\mathcal{I}_k}$, $\forall u(k-1) \in \mathbb{R}^{mM}$. From (25), it then follows that $\hat{x}(k) = x(k)$. Hence, (a) \Rightarrow (b) holds.

(b) \Rightarrow (a) : Suppose that (b) holds. To prove by contradiction, suppose that $|\mathcal{I}_k| \geq M/2$ and consider the non-active mode case $a_1(k) = 0$. Since $|\mathcal{I}_k| \geq M/2$, it follows from (27) and (28) that $\|v_{\tilde{T}}\|_1 \geq \|v_{\tilde{T}^c}\|_1$, $\forall v \in \ker(C_p) \setminus \{0\}$, i.e., C_p does *not* satisfy \tilde{T} -NSP. Thus, it follows from Theorem 1 that there exists $f_p(k) \in \tilde{\mathcal{F}}_{\mathcal{I}_k}$ such that $\hat{f}_p(k) \neq f_p(k)$, which means that $\hat{f}(k) \neq f(k)$ and the optimization problem in (18) does not provide a correct estimation. Indeed, this contradicts to the statement in (b) that $\hat{f}(k) = f(k)$, $\forall f(k) \in \mathcal{F}_{\mathcal{I}_k}$. Hence, (b) \Rightarrow (a) holds. \square

(Proof of Theorem 2): By Assumption 1, the leader is the active mode at $k = 0$ ($C = C_1$) and thus $\text{rank}(C_p) = \text{rank}(C) = nM$ at $k = 0$. This implies that the optimization problem in (18) yields $\hat{x}(0) = x(0)$, since $C = C_1$ has a trivial kernel and the solution of x satisfying $y(0) = Cx$ is uniquely determined. To provide the proof, suppose that (a) holds, i.e., $|\mathcal{I}_k| < M/2$ for all $k \in \mathbb{N} \setminus \{0\}$. Since $\hat{x}(0) = x(0)$, it follows from Lemma 1 that $\hat{x}(1) = x(1)$, $\hat{f}(1) = f(1)$, $\forall f(1) \in \mathcal{F}_{\mathcal{I}_1}$, $\forall u(0) \in \mathbb{R}^{mM}$. Since $\hat{x}(1) = x(1)$, it then follows that $\hat{x}(2) = x(2)$, $\hat{f}(2) = f(2)$, $\forall f(2) \in \mathcal{F}_{\mathcal{I}_2}$, $\forall u(1) \in \mathbb{R}^{mM}$. Similarly, it follows recursively by applying Lemma 1 that



Fig. 2. Simulation scenario. Each vehicle is able to measure relative state information with respect to the vehicle in front using on-board sensors (indicated as solid arrows).

$\hat{x}(k) = x(k)$, $\hat{f}(k) = f(k)$, $\forall f(k) \in \mathcal{F}_{\mathcal{I}_k}$, $\forall u(k-1) \in \mathbb{R}^{mM}$. Hence, (a) \Rightarrow (b) holds. Conversely, suppose that (b) holds. Since $\hat{x}(0) = x(0)$, it follows from Lemma 1 that $|\mathcal{I}_1| < M/2$. Similarly, it follows recursively by applying Lemma 1 that $|\mathcal{I}_k| < M/2$, $\forall k \in \mathbb{N} \setminus \{0\}$. Hence, (b) \Rightarrow (a) holds. \square

V. ILLUSTRATIVE EXAMPLE

In this section we provide an illustrative example to validate the proposed approach. To this end, consider a vehicle platoon that travels in a one-dimensional plane as illustrated in Fig. 2. Assume that there are $M = 9$ vehicles, where the lead vehicle (vehicle 1) is the leader and the others are the followers. Let $p_i \in \mathbb{R}$, $v_i \in \mathbb{R}$ be the position and the velocity of vehicle i , respectively. The dynamics of vehicle i ($i \in \{1, \dots, 9\}$) is given by

$$\dot{x}_i = \begin{bmatrix} 0 & 1 \\ 0 & -1/\tau \end{bmatrix} x_i + \begin{bmatrix} 0 \\ 1/\tau \end{bmatrix} u_i + f_i, \quad (29)$$

where $x_i = [p_i; v_i] \in \mathbb{R}^2$ and $u_i \in \mathbb{R}$ is the control input, $f_i \in \mathbb{R}^2$ is the fault signal, and $\tau = 0.2$ is the time constant. For simplicity, we assume that only velocity states are subject to faults, i.e., $f_i(k) = [0 \ f_{v,i}]^T$. Since we have $\dot{v}_i = -(1/\tau)v_i + (1/\tau)u_i + f_{v,i}$, one can also see that this situation is the case when actuators are subject to faults by physical faults or malicious attackers. We discretize the system in (29) in a zero-order hold fashion with $0.05s$ sampling time interval to obtain $x_i(k) = Ax_i(k-1) + Bu_i(k-1) + f_i(k)$. Each follower is able to measure the relative information with respect to the forwarding vehicle, i.e., $\{i+1, i\} \in \mathcal{E}$, $\forall i \in \{1, \dots, 8\}$. The control law for the leader is $u_1 = -v_1 + 1$ so that it moves with a constant speed $v = 1$. The control law for the followers is given by $u_i = -h(p_i - p_{i+1} - d)$ with $h > 0, d > 0$, $\forall i \in \{1, \dots, 10\}$, so that it keeps a constant distance with respect to the forwarding vehicle.

To indicate the active and non-active mode, we consider the following two situations; *outside the tunnel* and *inside the tunnel*. If the leader is outside the tunnel, it can utilize the GPS satellite to locate the state, which means that the leader is in active mode. On the other hand, if the leader is inside the tunnel, it cannot utilize the GPS due to signal loss by the presence of walls, which means that the leader is in non-active mode. The leader initially starts from outside of the tunnel to satisfy Assumption 1 and moves at the constant speed $v_1 = 1$, and enters the tunnel for a while. More specifically, we assume that

$$a_1(k) = \begin{cases} 0, & \text{if } k \in [150, 300], \\ 1, & \text{otherwise.} \end{cases} \quad (30)$$

Note that for the overall system derived in (12), the pair (A, C_1) is observable, while (A, C_0) is *not* observable. This

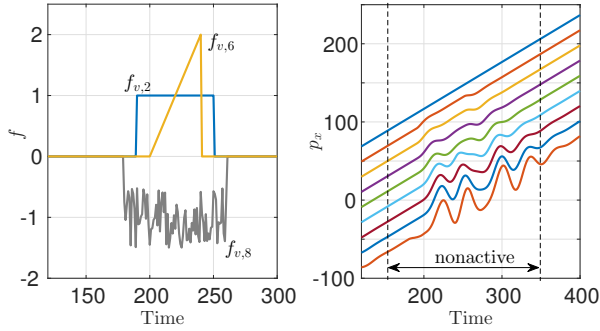


Fig. 3. True fault signals $f_{v,i}$ (left) and trajectories of p_i (right) for all vehicles.

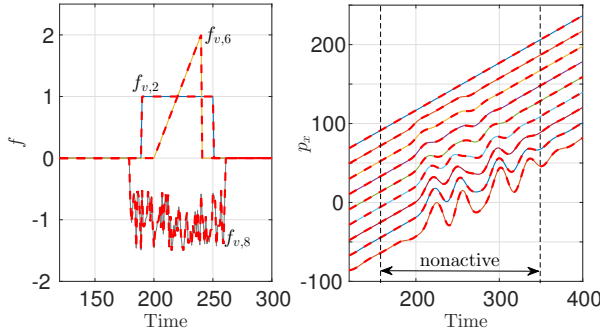


Fig. 4. Estimated faults (left) and states (right) by applying the proposed approach (18) (red dotted lines).

means that the observability is lost while the leader is non-active during $k \in [150, 300]$. Despite such un-observability, we will show below that the proposed estimator provides correct estimations for both states and fault signals. Fig. 3 illustrates fault signals that are injected to the vehicles and the actual trajectories of p_i . As shown in the figure, it is assumed that the vehicle 2, 6 and 8 are subject to faults with different shapes while the leader is in non-active mode. As a consequence, the resulting trajectories provide somewhat faulty behaviors as shown in Fig. 3. Fig. 4 illustrates the simulation results by applying the proposed estimator (red dotted lines). As shown in the figure, state and fault signals are appropriately estimated by applying the proposed approach. This is due to the fact that the number of faulty vehicles is given by $3 < M/2 = 4.5$, which satisfies the condition to guarantee correct estimations from Theorem 2.

Finally, Table I illustrates the cumulative estimation errors computed as $\sum_{k=1}^{350} \|x(k) - \hat{x}(k)\|_2$ against the number of faulty vehicles M_f . For each $M_f \in \{1, \dots, 7\}$, the set of faulty vehicles are selected as $\mathcal{I}_k = \{1, \dots, M_f\}$ for all $k \in [150, 200]$, and $f_{v,i}(k)$, $i \in \{1, \dots, M_f\}$ is assumed to be a random noise occurring from $f_{v,i}(k) \in [-10, 10]$. From the table, the estimation errors suddenly increase from $M_f = 5$. Indeed, this is exactly when the condition $\mathcal{I}_k < M/2 = 4.5$ is violated so that correct estimation is not given according to Theorem 2.

TABLE I
NUMBER OF FAULTY NODES AND THE RESULTING ESTIMATION ERRORS

M_f	1	2	3	4	5	6	7
	0.01	0.02	0.05	0.08	192	430	3811

VI. CONCLUSION

In this paper, we propose state and fault estimation strategies for multi-agent systems. The proposed estimator solves an ℓ_1 -norm optimization problem, which is based on the signal recovery in compressive sampling. We show that correct estimations are given if and only if the number of faulty nodes is less than half the total number of nodes. Finally, a numerical example of a platoon of vehicles illustrated the effectiveness of the proposed approach.

REFERENCES

- [1] M. Mesbhi and M. Egerstedt, *Graph Theoretic Methods in Multiagent Networks*, Princeton University Press, 2010.
- [2] Y. Arfat and F. E. Eassa, "A Survey on Fault Tolerant Multi-Agent System," *I.J. Information Technology and Computer Science*, vol. 9, pp. 39–48, 2016.
- [3] S. Ding, *Model-Based Fault Diagnosis Techniques: Design Schemes, Algorithms and Tools*, Springer, 2008.
- [4] I. Shames, A. Teixeira, H. Sandberg, and K. H. Johansson, "Distributed fault detection for interconnected second-order systems," *Automatica*, vol. 47, pp. 2757–2764, 2011.
- [5] C. Liu, B. Jiang, R. J. Patton, and K. Zhang, "Hierarchical Structure-based Fault Estimation and Fault-tolerant Control for Multi-agent Systems," *IEEE Transactions on Control of Network Systems*, 2018 (to appear).
- [6] Y. Li, H. Fang, J. Chen, and C. Yu, "Distributed Cooperative Fault Detection for Multiagent Systems: A Mixed H_∞/H_2 Optimization Approach," *IEEE Transactions on Industrial Electronics*, 2018 (to appear).
- [7] Y. Li, H. Fang, J. Chen, and C. Shang, "Distributed fault detection and isolation for multi-agent systems using relative information," in *Proceedings of the 2016 American Control Conference*, 2016, pp. 5939–5944.
- [8] D. Silvestre, P. Rosa, J. P. Hespanha, and C. Silvestre, "Distributed Fault Detection Using Relative Information in Linear Multi-Agent Networks," *IFAC-PapersOnLine*, vol. 48, no. 21, pp. 66–71, 2015.
- [9] P. P. Menon and C. Edwards, "Robust Fault Estimation Using Relative Information in Linear Multi-Agent Networks," *IEEE Transactions on Automatic Control*, vol. 59, no. 2, pp. 477–482, 2014.
- [10] E. J. Candès and M. B. Wakin, "An Introduction To Compressive Sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21–30, 2008.
- [11] S. Chen and D. Donoho, "Basis pursuit," in *Proceedings of the 28th Asilomar Conference on Signals, Systems and Computers*, 1994, pp. 41–44.
- [12] R. B. Bapat, *Graphs and Matrices*, Springer, 2010.
- [13] R. Gribonval and M. Nielsen, "Sparse representation in unions of bases," *IEEE Transactions on Information Theory*, vol. 49, no. 12, pp. 3320–3325, 2003.
- [14] D. L. Donoho and X. Huo, "Uncertainty principles and ideal atomic decomposition," *IEEE Transactions on Information Theory*, vol. 47, no. 7, pp. 2845–2862, 2001.
- [15] B. Efron, T. Hastie, and I. Johnstone, "Least angle regression," *Annals of Statistics*, vol. 32, no. 2, pp. 407–499, 2004.
- [16] H. Fawzi, P. Tabuada, and S. Diggavi, "Secure Estimation and Control for Cyber-Physical Systems under Adversarial Attacks," *IEEE Transactions on Automatic Control*, vol. 59, no. 6, pp. 1454–1467, 2014.
- [17] Y. Shoukry and P. Tabuada, "Event-Triggered State Observers for Sparse Sensor Noise/Attacks," *IEEE Transactions on Automatic Control*, vol. 61, pp. 2079–2091, 2016.
- [18] M. S. Chong, M. Wakaiki, and J. P. Hespanha, "Observability of linear systems under adversarial attacks," in *Proceedings of the 2015 American Control Conference*, 2015, pp. 2439–2444.