A Resilient Consensus protocol for networks with heterogeneous confidence and Byzantine Adversaries

David Angeli\textsuperscript{1} and Sabato Manfredi\textsuperscript{2}

\textbf{Abstract—} A class of Adversary Robust Consensus protocols is proposed and analysed. These are inherently non-linear, distributed, continuous-time algorithms for multi-agents systems seeking to agree on a common value of a shared variable, in the presence of faulty or malicious Byzantine agents, disregarding protocol rules and communicating arbitrary possibly differing values to neighbouring agents. We adopt monotone joint-agent interactions, a general mechanism for processing locally available information and allowing cross-comparisons between state-values of multiple agents simultaneously. The topological features of the network are abstracted as a Petri Net and convergence criteria for the resulting time evolutions formulated in terms of suitable structural properties of its invariants (so called siphons). Finally, simulation results and examples/counterexamples are discussed.

\textbf{Index Terms—} Network analysis and control, Cooperative control, Distributed control, Fault tolerant systems, Petri Nets

I. INTRODUCTION AND MOTIVATIONS

During recent years consensus problems have been studied in several situations including higher order agents dynamics, linear/nonlinear time varying Multi Agent Systems, symmetric and asymmetric agents’ interactions [6], [7], [13] (just to cite a few). Many fundamental problems of modern Cyber-Physical Systems can be reduced to an underlying consensus framework for Multi Agent Systems, including for instance distributed optimization, formation control (e.g. the multi-robot rendezvous), synchronization of the electric power grid, sensor fusion, distributed estimation and control (see i.e. [14]–[17] and references therein). A major concern in such applications of networked systems is to guarantee security at the application layer protocols, viz. with respect to malicious agents able to attack the Cyber layer, but unable to tamper with the Physical layer. This has been a major driving force to formulate resilient consensus algorithms such that, in the event of an attack in which some nodes are compromised, the remaining (healthy) nodes are still able to achieve their objective (or at least a relaxed version of the objective). Within this line of investigation, the problem of Adversary Robust Consensus Protocols (ARC-P) was formulated in [10], following earlier seminal results in [21]. Therein, Leblanc and coworkers propose and analyze a protocol which allows \( n \) cooperating agents to converge towards a consensus state, within a complete all-to-all network, even when a subset of agents (of cardinality up to \( \lfloor n/2 \rfloor \)) is malicious or faulty, namely it evolves in a completely arbitrary way, with the sole constraint of broadcasting its own state to all remaining agents. The proposed protocol simply orders state values in ascending (or descending order) and removes \( F \) top and lowest values from the ordered list, where \( F \) is an apriori fixed bound to the number of malicious agents. Then, the average among the remaining values is computed and a standard linear consensus update equation is applied.

Among various types of fault scenarios, malicious agents which can convey different state trajectories to different neighbours in the network, are those of major concerns and interest in Cyber-Physical Systems. Agents with this ability are usually referred to as Byzantine, and Byzantine consensus protocols exhibit robustness to such kind of threats. In this case the main drawback is that the ability of malicious agents of differentiating the information sent to neighbors may disrupt consensus even when adversarially robust consensusability is fulfilled as previously discussed.

In this respect subsequent analysis has been devoted in [12] to the important topic of relaxing the all-to-all topology requirement and investigating sufficient conditions for Adversary Robust Consensus on the basis of local information only, or in the presence of Byzantine agents [11], who may, either intentionally or due to faulty conditions, communicate different state values to different neighbors. A related line of investigation assumes the presence of trusted nodes, [3], or analyses group consensus in the presence of structured and unstructured Byzantine faults [4].

The main contribution of this paper can be summarised as follows. First, we broaden the applicability of joint-agent interactions originally introduced to improve resilience for CPS with malicious agents ( [1], [2]) to the case in which Byzantine agents are present. Such formalism is very flexible and encompasses most of scenarios considered in the literature (i.e. [8], [9], [12]), allowing to partition neighbours of every agent in several subgroups, to be suitably sorted, reduced and averaged while adding (possibly with different weights) the influences resulting from distinct subgroups as a final step (e.g. heterogeneous confidence). This type of local information processing results in consensus protocols which allow different levels of trust or reputation attributed to different set of neighbors and, generally speaking, break the symmetry implicit in the use of a single sorting and reducing function.

\textsuperscript{1} David Angeli is with the Control and Power Group, Electrical and Electronic Engineering Department, Imperial College, London and Dip. di Ingegneria dell’Informazione, University of Florence, Italy. Email: d.angeli@imperial.ac.uk

\textsuperscript{2} Sabato Manfredi is with the Department of Electrical Engineering and Information Technology, University of Naples Federico II, Italy. Email: sabato.manfredi@unina.it
or in the topological assumption of treating as equals all neighbours (arising from a description of agents interaction by means of standard graphs). Such feature plays a focal role for dealing with the safety of modern Cyber-physical systems where individual and group are characterised by different level of trust/reputation ( [20]).

Secondly, an extended class of intrinsically nonlinear consensus protocols is presented and tight sufficient conditions for consensus in the presence of Byzantine agents are provided. Such protocols can be characterized, from a topological point of view, as bipartite graphs, and more specifically Petri Nets. The results show that number and connectivity of healthy and malicious agents are crucially interlinked in ensuring consensus. The question of which topology is the best in order to guarantee robustness and what is the maximum number of Byzantine agents that can be allowed for a given topology is of great interest, but outside the scope of the current paper.

This is, to the best of our knowledge, the first attempt to highlight in a systematic fashion by the formalism of Petri Nets the possibility of retaining asymptotic convergence properties in the face of arbitrary exogenous disturbances as determined by Byzantine agents. It is worth pointing out that the result assumes such agents are not updating their own state according to the consensus protocol rather in arbitrary continuous evolutions, and intentionally communicate different evolutions to different neighbors. Finally, the structural notions, developed in the context of Petri Nets to ascertain their liveness as Discrete Event Systems, play a crucial role in characterizing the ability of a network of agents to reach consensus regardless of initial conditions.

II. PROBLEM FORMULATION

This manuscript derives tight characterizations of when agents implementing a consensus protocol through joint-agent interactions, may be able to exhibit resilience in the presence of possibly malicious Byzantine agents. In particular, we consider the following class of nonlinear finite-dimensional network’s dynamics:

\[ \dot{x} = f(x) \]  

with state \( x \in \mathbb{R}^n \), and \( f : \mathbb{R}^n \rightarrow \mathbb{R}^n \) a Lipschitz continuous function, describing the update laws of a finite set of agents \( \mathcal{N} = \{1, 2, \ldots, n\} \) on the basis of their own and neighbours’ state values. For the sake of simplicity, we assume \( f_j \) to be monotonically non-decreasing with respect to all \( x_i \) (\( i \neq j \)) so that the resulting flow is monotone with respect to initial conditions when the standard component-wise partial order between vectors is considered, citemonotone. A significant amount of literature, in recent years, has focused on conditions under which solutions of (1) asymptotically converge towards consensus:

\[ \lim_{t \rightarrow +\infty} \varphi(t, x_0) = \tilde{x} \mathbf{1} \]

for some \( \tilde{x} \in \text{co}(\{x_i(0), i \in \mathcal{N}\}) \), where \( \mathbf{1} \) is the vector of all ones (of compatible dimension), \( \text{co}(\cdot) \) denotes the convex hull and \( \varphi(t, x_0) \) is the solution of (1) at time \( t \) from initial condition \( x_0 \).

In this paper we consider a more challenging scenario in which agents are partitioned according to \( \mathcal{N} = \mathcal{H} \cup \mathcal{B} \), respectively the set of Healthy and Byzantine agents respectively. Without loss of generality, the state vector \( x \) is partitioned into two subvectors, \( x_H \) and \( x_B \) and \( f(x) \) as \( [f_H(x)' , f_B(x)']' \).

Our aim is to study conditions on

\[ \dot{x}_H^i = f_H(x_H, x_B) \]

for \( i \in \mathcal{H} \) such that solutions of (3) asymptotically converge to equilibria of the type:

\[ \lim_{t \rightarrow +\infty} x_H(t) = \tilde{x} \mathbf{1}_H \]

for all initial conditions \( x_H(0) \) and all exogenous input signals \( x_B^i(\cdot) \). Notice that equation (3) is expressed in component-wise form, with respect to index \( i \in \mathcal{H} \), to emphasize that Byzantine agents are free to send different signals to every neighbour \( i \in \mathcal{H} \), which arranged in a vector are denoted as \( x_B^i \). A more formal definition follows.

Definition 1: We say that network (1) achieves robust consensus in the face Byzantine malicious agents in \( \mathcal{B} \), if for all \( x_H(0) \), and all uniformly bounded exogenous input signals \( x_B^i(\cdot), i \in \mathcal{H}, \) there exists \( \tilde{x} \in \text{co}(\{x_i(0), i \in \mathcal{N}\}) \) such that:

\[ \lim_{t \rightarrow +\infty} \varphi_H(t, x_H(0), x_B^i(\cdot)) = \tilde{x} \mathbf{1}_H, \]

(4)

where \( \varphi_H(t, x_H(0), x_B^i(\cdot)) \) denotes the solution of (3) at time \( t \) from initial condition \( x_H(0) \) and input signals \( x_B^i(\cdot), i \in \mathcal{H} \). In applications, one would typically seek resilience with respect to several possible combinations of Byzantine agents (corresponding to different choices of set \( \mathcal{B} \) and, for each one of them, verify conditions for asymptotic convergence towards consensus of the remaining healthy agents \( \mathcal{H} \).

The consensus value reached in the presence of malicious agents is not necessarily the same as the one achieved when all agents are healthy. In fact, even a malicious agent can influence the final consensus value. However, under suitable conditions investigated next, it is unable to prevent consensus of healthy agents. We formulate the requirements on the flow of information and influence needed for achieving such kind of behaviour, by adopting the notion of joint agent interaction, [1]. To this end let \( e_i \) denote a vector in \( \{0, 1\}^d \), of dimension \( d \) implicit from the context and support equal to \( I \subset \mathcal{N} \).

Definition 2: We say that a group of agents \( I \subset \mathcal{N} \) jointly influence agent \( j \in \mathcal{N} \setminus I \) if for all compact intervals \( K \subset \mathbb{R} \) there exists a positive definite function \(^1 \rho \), such that, for all scalar \( x_1, x_j \in K \) it holds:

\[ \text{sign}(x_1 - x_j)f_j(x_j e_{\mathcal{N} \setminus I} + x_1 e_i) \geq \rho(|x_1 - x_j|). \]

We denote this by the following shorthand notation: \( I \rightarrow j \). Intuitively, Definition 2 quantifies how the simultaneous displacement of agents in \( I \) from a consensus value \( x_I \) towards \( x_I \) affects the time derivative of agent \( j \). Notice that the relation \( I \rightarrow j \), is monotone (in \( I \)) with respect to set-inclusion, viz. for any \( j \in \mathcal{N} \setminus I \) we have:

\[ I \rightarrow j \text{ and } I \subset I \Rightarrow I \rightarrow j. \]

\(^1\)A function \( \rho : [0, +\infty) \rightarrow [0, +\infty) \) is positive definite if it is continuous, \( \rho(0) = 0 \) and \( \rho(r) > 0 \) for all \( r > 0 \).
For this reason, it is normally enough to consider minimal influences alone. We say that $I$ influences $j$ and that this influence is minimal if there is no $I \subseteq I$ such that $I \rightarrow j$.

Remark 1: Equation (4) plays a crucial role in allowing to characterize the asymptotic behaviour of the agents and the link between dynamics and its abstraction using Petri Nets. The rationale behind it is that for large times the maximum (and minimum) of agents’ states will always settle to some constant values, and their time derivatives to 0. By virtue of (4), agents achieving the maximum (or the minimum), will only be influenced by groups that include at least a Byzantine agent or another healthy agent achieving the maximum (the minimum, respectively). This introduces non-trivial constraints in the way that the maximum or the minimum may be achieved, and (under suitable topological conditions later discussed) allows to conclude that the maximum must coincide with the minimum.

III. PETRI NET REPRESENTATION

To achieve a graph-theoretical characterization of the network of interactions which afford resilience towards a given set of Byzantine agents, we need an effective representation of joint-agent interactions. We adopt, to this end, the formalism introduced in [1] and borrow the language of Petri Nets to represent multi-agent systems. Petri Nets (usually a Discrete Event Systems’ abstraction) are hereby used only as bipartite graphs, so as to capture interconnection topology. An (ordinary) Petri Net is a quadruple $\{P, T, E_I, E_O\}$, where $P$ and $T$ are the (finite) sets of places places and transitions, respectively. These are nodes of a directed bipartite graph. Edges are also of two types: $E_I \subset T \times P$ connecting transitions to places and $E_O \subset P \times T$ connecting places to transitions. In our context places represent agents while transitions stand for interactions among them. Specifically, each agent $i \in N$ is represented by a unique associated place $p_i \in P$. Moreover, if agents in $J \subset N$ jointly influence agent $i$, this is denoted as $J \rightarrow i$ and, provided this interaction is minimal, it is represented by a single transition $t \in T$, with edges $(p_j, t) \in E_O$ for all $j \in J$ and a single edge $(t, p_i)$ in $E_I$. By construction every transition can be assumed to afford exactly one outgoing edge, (unlike in general Petri Nets).

As for standard non-resilient protocols, the topology of interconnections is crucial in order to characterize networks with the ability to guarantee asymptotic emergence of consensus. The following notions will be central in carrying out such characterization. The input transitions for a set of places $S \subset P$, are denoted as

$$I(S) = \{t \in T : \exists p \in S : (t, p) \in E_I\},$$

and, its output transitions are:

$$O(S) = \{t \in T : \exists p \in S : (p, t) \in E_O\}.$$

Definition 3: A non-empty set $S \subset P$ is called a siphon if $I(S) \subseteq O(S)$. A siphon is minimal if no proper subset is also a siphon.

Remarkably, in a group of agents that correspond to a siphon, any influence acting upon elements of $S$ needs to come (at least in part) from within the group. Computing siphons is in general an NP-complete combinatorial problem, [19], however efficient algorithms exist for relevant classes of Petri Nets, (see [18]). Authors of [1] show that the ability to asymptotically converge towards consensus (regardless of initial conditions) is related to the siphon overlapping property. A related condition can be formulated for adversarially robust consensus, when faulty or malicious agents are not Byzantine, viz. they broadcast the same state value to all of their neighbours. The following is the appropriate notion for adversarially robust consensus:

Definition 4: A non-empty set $S \subset P$ is an $F$-controlled siphon (with $F \subset P$ and $F \cap S = \emptyset$), if:

$$I(S) \subset O(S) \cup O(F).$$

We call the set $F$ the switch of siphon $S$.

Notice that Definition 4 boils down to the standard notion of siphon for $F = \emptyset$. Union of $F$-controlled siphons is again an $F$-controlled siphon and, in particular, if a set is an $F_1$-controlled siphon it is also an $F_2$-controlled siphon for all $F_2 \supseteq F_1$. The intuition behind this Definition is that malicious agents in $F$ may prevent healthy agents in $S$ from increasing (or decreasing). This is achievable by communicating corrupted state values which are either below the minimum or, respectively, above the maximum of all values within the siphon. For the sake of completeness we recall the notion of siphon-switch overlapping property which is appropriate for Adversary Robust Consensus for non Byzantine agents, [2].

Definition 5: We say that a Petri Net fulfils the siphon/switch overlapping property with respect to faults in $F \subset N$ if $H := N \setminus F$ is a siphon and for all pairs of controlled siphons $S_1, S_2 \subset H$ and associated switches $F_1, F_2 \subset F$, we have the following:

$$S_1 \cap S_2 = \emptyset \Rightarrow F_1 \cap F_2 \neq \emptyset. \quad (5)$$

Remarkably the simpler siphon-overlapping property is the appropriate concept for Byzantine agents.

Definition 6: We say that a Petri Net fulfils the siphon overlapping property with respect to Byzantine agents in $B \subset N$ if $H = N \setminus B$ is a siphon of the Petri Net and for all pairs of controlled siphons $S_1, S_2 \subset H$ and associated switches $F_1, F_2 \subset B$, we have the following:

$$S_1 \cap S_2 \neq \emptyset. \quad (6)$$

We remark that while condition (6) is simpler to state than (5), its fulfilment is a stronger requirement. Consider, as the simplest possible example, the all-to-all network of joint-agent interactions associated to $N = \{1, 2, 3\}$. Assume $B = \{3\}$ and $H = \{1, 2\}$. It is easy to see that $H$ is a siphon and that $\{1\}$ and $\{2\}$ are the only $B$ controlled siphons. Hence, the net does not fulfill Definition 6, since $\{1\} \cap \{2\} = \emptyset$. On the other hand, it fulfills Definition 5, since despite the empty intersection we have $F_1 = F_2 = F_1 \cap F_2 = B \neq \emptyset$. Intuitively, a Byzantine agent is able to disrupt consensus in the considered Net, since it might communicate different values to agents 1 and 2, for instance the value $x_1(t)$ to agent 1 and $x_2(t)$ to agent 2, so that neither one will move from their own current value. On the other hand, an agent broadcasting a
single malicious value to both agents will always end up either attracting towards lower values the agent with maximum state value, or attracting towards higher values the agent with the minimum state value. Such a situation occurs because the two siphons have overlapping switches.

IV. MAIN RESULT AND SKETCH OF PROOFS

In the following Section we state the main result and clarify the steps of its proof. We consider equations of multi-agent systems derived on the basis of the associated Petri Net, so as to establish a close link between the considered dynamics and the Petri Net structural properties. To fix ideas, for any \( I \subset P \) \( \alpha \) and \( j \notin I \), we consider the function:

\[
f_{I \rightarrow j}(x) := \max_{i \in I} \min \{x_i - x_j, 0\} + \min \max \{x_i - x_j, 0\}.
\]

(7)

Equation (7) defines the simplest form of joint-agent interaction, computing (through \( \max \{x_i \} \) and \( \min \{x_i \} \)) the signed point-to-set distance between agent \( j \) and the convex-hull of state values for agents in \( I \). Given any transition \( t \in T \), denote by:

\[
I(t) := \{p \in P : (p, t) \in E_O\}
\]

and by \( j(t) \) the unique place such that \((t, j(t)) \) belongs to \( E_I \). For a given Petri Net \( \{P, T, E_I, E_O\} \), we could therefore consider the following set of equations:

\[
\dot{x} = \sum_{t \in T} \alpha_t f_{I(t) \rightarrow j(t)}(x) e_{j(t)}
\]

(8)

for any choice of constant coefficients \( \alpha_t > 0 \) and recalling that \( e_j \) denotes the \( j \)-th element of the canonical basis of \( \mathbb{R}^n \). More general ways of defining systems equations are also possible, while still fulfilling the requirements of joint-agent interactions introduced in Section II. These are outlined in [2]. Notice that computation of (8) can be carried out in a distributed way provided agent \( j \) is aware of all state values of neighbouring agents, defined as:

\[
N_j := \{k \in N : \exists I \subset N : k \in I \text{ and } I \rightarrow j\}.
\]

We are now ready to formulate the main result and to illustrate the technical steps of its derivation.

Theorem 1: Consider a cooperative network of agents \( N \) as in (8) and let \( N \) be the Petri Net associated to its set of minimal joint agent interactions. Consider a partition of \( N \) into two disjoint subgroups \( B, \mathcal{H} \), the Byzantine and Healthy agents (respectively), along with the projected dynamics, (3). Then, asymptotic consensus is achieved among the agents in \( \mathcal{H} \) provided \( N \) fulfills the siphon overlapping property with respect to Byzantine agents in \( B \).

It is worth pointing out that the result assumes Byzantine agents are communicating to all neighbouring agents arbitrary and potentially different continuous (bounded) evolutions. In order to prove the result, let us generalize a useful Lemma from [2].

Lemma 1: Assume \( \mathcal{H} \) is a siphon of net \( N \). Consider equations (8) and denote with \( x_H \) the state vector of agents in \( \mathcal{H} \), along with the corresponding dynamics

\[
\dot{x}_H(t) = f_H(x_H(t), x_B(t))
\]

(9)
as introduced in (3), with \( i \in \mathcal{H} \). Then, for any \( c \in \mathbb{R} \), the sets:

\[
\mathcal{F}_c := \{x_H \in \mathbb{R}^{|H|} : x_H \leq c I_H\}, \\
\mathcal{F}_c := \{x_H \in \mathbb{R}^{|H|} : x_H \geq c I_H\}
\]

are robustly forward invariant for any bounded input signals \( x_B(\cdot) \), \( i \in \mathcal{H} \).

Proof: Let \( x_H(t) \in \partial \mathcal{F}_c \) (the boundary of \( \mathcal{F}_c \) ) be arbitrary, and \( h \in \mathcal{H} \) be any agent with \( x_h = c \). To show invariance of \( \mathcal{F}_c \) we need \( f_H(x, x_B) \leq 0 \) for all \( x_B \in K \). This condition, in fact, amounts to \( f_H(t, x, x_B) \in TC_{x, x_B}(\mathcal{F}_c) \) for all \( x \in \partial \mathcal{F}_c \) and all \( x_B \in K \).\( \forall h \in \mathcal{H} \) (where \( TC_{x, x_B}(\mathcal{F}_c) \) denotes the Bouligand's tangent cone of \( \mathcal{F}_c \) at \( x_H(t) \)). This, in turn, implies forward invariance of \( \mathcal{F}_c \) by Nagumo's Theorem, [22]. To show the above inequality let \( O(p_h) = \{t_1, \ldots, t_{|O(p_h)|}\} \). Since \( \mathcal{H} \) is a siphon, for all \( t_i \) in \( O(p_h) \) there exists \( h \in I(t_i) \cap \mathcal{H} \). Hence, recalling definition (7),

\[
\dot{f}_{I(t_i) \rightarrow h}(x) \leq f_{I(t_i) \rightarrow h}(ce_H, x_B) = c - c = 0.
\]

Exploiting the expression for \( \dot{x}_h \), yields

\[
\dot{x}_h = \sum_{t \in T \cap I(t_i) \rightarrow h} \alpha_t f_{I(t_i) \rightarrow h}(x) \leq 0.
\]

A symmetric argument works for \( \mathcal{F}_c \). This completes the proof of the Lemma.

In the following we sketch the main idea of the proof of Theorem 1.

Proof: Let \( x_B(t) = [\ldots x_B^t(t) \ldots] \), for \( h \in \mathcal{H} \) be an arbitrary bounded, continuous signal. Specifically, take \( x_B(t) \in K^{\text{card}(\mathcal{H})} \) for some compact set \( K \subset \mathbb{R}^{\text{card}(B)} \). For an arbitrary initial agent configuration \( x_H(0) \) define the evolution of healthy agents according to equation (3). In particular, denote the solution \( x_H(t) := \varphi_{H}(t, x_H(0), x_B(\cdot)) \), for all \( t \geq 0 \). Moreover, we let:

\[
\tau_{H} := \max_{h \in \mathcal{H}} x_h, \quad \tau_{H} := \min_{h \in \mathcal{H}} x_h.
\]

Since \( \mathcal{H} \) is a siphon, by Lemma 1, we see that for all \( \tau_2 \geq \tau_1 \geq 0 \):

\[
x_H(t_1) \in \mathcal{F}_H(t_1) \Rightarrow x_H(t_2) \in \mathcal{F}_H(t_2).
\]

In particular then, \( \tau_{H}(t_1) \geq \tau_{H}(t_2) \), viz. \( \tau_{H} \) is monotonically non-increasing. A symmetric argument applies to \( x_B(t) \), which is monotonically non-decreasing. Henceforth, \( x_B(t) \) is uniformly bounded and the limits

\[
\tau_{H} := \lim_{\tau \rightarrow +\infty} \tau_{H}(\tau), \quad \tau_{H} := \lim_{\tau \rightarrow +\infty} \tau_{H}(\tau),
\]

(10)

exist finite. To study the asymptotic dynamics of the system it is useful to define the convex-valued differential inclusion given below:

\[
\dot{z} \in F_H(z) := \co \bigg( \bigcup_{x_B^t \in K^{\text{card}(\mathcal{H})}} \{f_H(z, x_B^t)\} \bigg).
\]

(11)

Due to compactness of \( K \), and Lipschitz continuity of \( f_H \), \( F_H \) is a Lipschitz continuous set-valued map. In particular,
$x_H(t)$, is also a (bounded) solution of (11). Consider next the associated \( \omega \)-limit set, which, by boundedness of $x_H(t)$, is non-empty and compact:

$$
\Omega_H := \left\{ x : \exists \{\tau_n\}_{n=1}^{+\infty} : \lim_{n \to +\infty} \tau_n = +\infty \right\} \qquad (12)
$$

and $x = \lim_{n \to +\infty} x_H(\tau_n)$.

Notice that, by definition, for any $z_H \in \Omega_H$ we have $\tau_H = \tau_H^\infty$ and $\check{z}_H = \check{z}_H^\infty$. As is well known, $\Omega_H$ is a weakly invariant set for the differential inclusion (11). Selecting any element $z_H$ in $\Omega_H$, there exists at least one viable solution $z_H(t)$ of (11), such that $z_H(t) \in \Omega_H$, for all $t$. Notice that, by Lipschitzness of $F_H$, the sets

$$
M(t) := \{ h \in H : z_h(t) = \tau_H \}
$$

and

$$
m(t) := \{ h \in H : z_h(t) = \check{z}_H \}
$$

are monotonically non-increasing with respect to set-inclusion and, trivially, non-empty for all $t \geq 0$. Hence, there exists some finite $t \geq 0$ such that $M(t) = M(\tau)$ and $m(t) = m(\tau)$ for all $t \geq \tau$. Following the same steps as in [2] it is possible to show that, there exist $F_M, F_m \in \mathcal{B}$ (possibly empty) such that $M(\tau)$ is an $F_M$ controlled siphon and $m(\tau)$ is an $F_m$ controlled siphon. Hence, by the siphon overlapping property we have: $M(\tau) \cap m(\tau) \neq \emptyset$ which in turn implies $\tau_H^\infty = \check{z}_H^\infty$, thus proving robust asymptotic consensus.

V. Examples and simulations

We consider a set of $n \geq 5$ agents arranged in a ring topology. In particular, $\mathcal{N} = \{1, 2, \ldots, n\}$ and we assume for each $k \in \mathcal{N}$, a set of neighbours $\mathcal{N}_k$ defined as follows:

$$
\mathcal{N}_k = \{ m \neq k \in \mathcal{N} : \min\{|m-k|, |m+n-k|\} \leq 2 \},
$$

which corresponds to a 4-Regular topology with each agent having exactly 4 neighbours. For any agent $k \in \mathcal{N}$, we assume the following set of joint-agent interactions: $\{ I \to k : I \in \mathcal{N}_k : |I| = 2 \}$, viz. every agent is influenced by any possible pair of its neighbours. A set $\Sigma$ is a siphon, if whenever $k$ belongs to $\Sigma$, any influence on $k$ comes (at least in part from the siphon). This amounts to requiring that for every element of $\Sigma$ at most one of its neighbours is not in $\Sigma$. This condition is equivalent to asking that elements of $\mathcal{N} \setminus \Sigma$ are at least 4 agents apart, viz. for all $m_1 \neq m_2 \in \mathcal{N} \setminus \Sigma$, it holds $\min\{|m_1-m_2|, |m_1+n-m_2|\} \geq 4$. Notice that, siphons are very big with few “gaps” along the ring, and therefore they clearly fulfill the siphon overlapping property in case of all agents being healthy. Let us consider next the case of a single malicious Byzantine agent, which we assume to be agent 1, without loss of generality. Let $F = \{1\}$. A set $\Sigma$ is an $F$-controlled siphon if and only if for all $m_1 \neq m_2 \in \mathcal{N} \setminus (\Sigma \cup F)$, it holds $\min\{|m_1-m_2|, |m_1+n-m_2|\} > 4$. In other words, $\Sigma$ is a siphon if its gaps are at least 4 agents apart from each other and from agent $1$. In particular, $\mathcal{N} \setminus \{1\}$ is clearly a siphon and the siphon overlapping property still holds, thus proving resilience with respect to a single Byzantine agent. If, on the other hand, 2 Byzantine agents are considered next to each other, for instance agents $\{1, 2\}$, it is immediate to see that $\mathcal{N} \setminus \{1, 2\}$ is not a siphon (due to the presence of 2 gaps next to each other), and therefore resilience to these malicious agents cannot be ensured. To understand why, consider the case in which all other agents, $\{3, \ldots, n\}$ have already reached agreement (for instance at 0). Then agents 1 and 2 may broadcast a value of +1 to agent 3, and a value of −1 at agent $n$, thus inducing agent 3 to grow and agent $n$ to decrease. This proves that the consensus condition is not robust when dealing with 2 Byzantine agents sitting next to each other.

The effectiveness of the proposed conditions is validated by numerical simulation considering a 4-Regular network of 7 nodes with the Byzantine node 7 generating malicious signals $x_7^j$ to nodes $j = 1, 2, 5, 6$: $x_7^1(t) = 5 + \frac{4}{10} \sin\left(\frac{t}{40}\right)$, $x_7^2(t) = 15 + \frac{4}{10} \sin\left(\frac{t}{40}\right)$, $x_7^5(t) = 10 \cos\left(\frac{t}{40}\right)$, $x_7^6(t) = -\frac{4}{10} \sin\left(\frac{8t}{40}\right) \sin\left(\frac{t}{40}\right)$.

Fig. 1 confirms the asymptotic convergence of the state of healthy agents (i.e. $x_i(t)$, $i = 1, \ldots, 6$) to the consensus equilibrium. It is worth pointing out that while the “faulty” signals are apriori generated in an open-loop fashion, similar results can be achieved even under the assumption of malicious signals generated in closed-loop by agent 7. To this end we consider the scenario in which the initial condition $x(0)$ fulfills the inequalities, $x_1(0) > x_2(0) > x_3(0) > x_4(0) > x_5(0) > x_6(0)$ and the Byzantine agent 7 communicates the values achieved according to the following feedback strategy:

$$
x_7^1(t) = x_7^2(t) = \max\{x_1(t), x_2(t)\}
$$

$$
x_7^5(t) = x_7^6(t) = \min\{x_5(t), x_6(t)\}.
$$

This is, in some sense the most disruptive strategy for agent 7, as it sends high values to agents with the highest state values, and low values to those with lowest state values, in order to prevent them coming closer together. However, as shown in simulation and confirmed by the theory, the agent is unable to prevent asymptotic consensus, (see Fig. 2).

We show next how a related all-to-all protocol does not exhibit resilience to Byzantine agents. Specifically, considered
Byzantine agents. The criteria avail the possibility of testing, invariants, of the ability of nonlinear consensus protocols with through the language of Petri Nets and associated structural for all $t \geq 0$ assuming that $x_{\mathcal{S}}(0) = x_{\mathcal{S}}$. This protocol, simply computes the median of all agents and updates agents so as to converge asymptotically to it. From the point of view of the associated Petri Net this corresponds to the following set of interactions:

$$\{ I \rightarrow j : \text{card}(I) = 3 \text{ and } j \notin I \}. \quad (14)$$

Notice that each agent is influenced by all possible triplets of neighbouring agents. There are 4 such triplets for each agent. This is because in order to shift the value of the median of a vector of all ones of dimension 5 (in order to fulfil Definition 4), one needs to increase (or decrease) at least 3 of its entries. A set $S \subset \mathcal{N}$ is a siphon for the Petri Net associated to (14) if and only if card($S$) $\geq 3$. Consider next, without loss of generality, the case of $\mathcal{B} = \{3\}$ and $\mathcal{H} = \{1, 2, 3, 4\}$. A set $S \subset \mathcal{H}$ is a $\mathcal{B}$-controlled siphon if and only if card($S$) $\geq 2$. In particular, then $S_1 := \{1, 2\}$ is a $\mathcal{B}$-controlled siphon, and $S_2 := \{3, 4\}$ is yet another $\mathcal{B}$-controlled siphon. Clearly this violates the siphon overlapping property as $S_1 \cap S_2 = \emptyset$. Hence the network (13) does not guarantee consensus in the presence of a single Byzantine agent (it does however guarantee consensus in the presence of up to two non Byzantine malicious agents). This is easily verified by assuming $x_1(0) = x_2(0) = 0$ and $x_3(0) = x_4(0) = 1$ and assuming that $x_3(t) \equiv x_4(t) = 0$ and $x_5(t) \equiv x_6(t) = 1$ for all $t \geq 0$. Such a choice will result in solutions fulfilling $x_{\mathcal{H}}(t) = x_{\mathcal{S}}(0)$ for all $t \geq 0$, hence disrupting the on-set of asymptotic consensus among healthy agents.

**VI. CONCLUSIONS**

The paper presents a topological characterization, expressed through the language of Petri Nets and associated structural invariants, of the ability of nonlinear consensus protocols with joint-agent interactions to achieve resilience with respect to Byzantine agents. The criteria avail the possibility of testing, when a given set of agents are malicious (and/or faulty) and Byzantine, whether the complementary set of agents retain the ability to achieve asymptotic consensus, for arbitrary initial conditions. This approach generalizes the class of resilient protocols that may be analysed and formulated to encompass monotone joint-agent interactions. Simulation results and examples are discussed. The paper paves the way to future studies into the problem of selecting specific topologies of interactions in order to maximize the resilience achievable subject to assigned topological sparsity constraints.

**REFERENCES**


