Collaborative Vehicle Platoons with Guaranteed Safety Against Cyber-Attacks

Keijzer, Twan; Chanfreut, Paula; Maestre, José María; Ferrari, Riccardo Maria Giorgio

DOI
10.1109/TITS.2024.3503370

Publication date
2025

Document Version
Final published version

Published in
IEEE Transactions on Intelligent Transportation Systems

Collaborative Vehicle Platoons With Guaranteed Safety Against Cyber-Attacks

Twan Keijzer, Paula Chanfreut, José María Maestre, *Senior Member, IEEE*, and Riccardo Maria Giorgio Ferrari, *Member, IEEE*

***Abstract*— The wireless communication used by vehicles in collaborative vehicle platoons is vulnerable to cyber-attacks, which threaten their safe operation. To address this issue we propose a topology-switching coalitional *model predictive control* (MPC) method based on a reduced order unknown input observer which detects and isolates the cyber-attacks, so that the attacked communication links can be disabled by means of a topology switch. Also, the MPC controller is designed to guarantee robustness against undetected attacks and the increase of uncertainty derived from disabling communication links. The proposed control method also conforms to a relaxed string stability condition and is guaranteed to be safe against crashes.**

***Index Terms*— Cyber-attack tolerant control, coalitional control, topology-switching control, model predictive control.**

I. Introduction

AUTONOMOUS and collaborative vehicles are being broadly researched as solutions to problems like road congestion, pollution, and accidents caused by human error. One of the solution concepts that has been proposed is that of *collaborative vehicle platoons* (CVP) following a lead vehicle [1], [2], since vehicles in a CVP can drive more consistently and closer together than human-driven vehicles, thus reducing pollution, congestion, and road accidents.

In the last decades, several control methods have been proposed for platooning problems [3], [4], including consensus-based approaches [5], robust strategies [6], [7], and model predictive control (MPC) [8], [9]. In this article, we focus on *distributed model predictive control* (DMPC) [10], i.e. strategies where a set of local MPC agents communicate to coordinate their actions. Within this framework, this article follows the *coalitional* strategy [11], [12], [13], [14], which is characterized by the dynamic formation of disjoint clusters of cooperative agents, the so-called *coalitions*. See [15], [16], [17] for examples of its application in irrigation canals, traffic systems, and solar parabolic plants. In each coalition, local MPC agents share data to attain their control goals, whereas the cooperation between coalitions is reduced to a minimum. Therefore, the coordination efforts are reduced in comparison to classical cooperative DMPC approaches. As discussed in [14], this inherent capacity to handle multiple communication topologies makes this type of controllers a suitable and scalable solution to deal with communication failures.

In this regard, [18], [19], [20], [21], [22] deal with CVPs where the communication topology switches due to vehicles joining and leaving the platoon, the possible inter-vehicles communication failures, and the existence of a minimum distance to communicate. These works stress the relevance of flexible controllers able to accommodate these dynamic communication constraints while providing performance and stability guarantees. In particular, by using the results of [8], the work of [18] presents a DMPC for platoons with switching topologies that guarantees convergence of the predicted terminal states. Additionally, [22] proposes a switching control law to achieve string stability in heterogeneous platoons with communication losses, and [23] studies the influence of the communication topology on the stability and scalability of platoons considering linear feedback controllers.

The exchange of data in the CVP can also be subject to cyber-attacks, which threatens its safe operation [24], [25], [26]. Therefore, controllers able to detect and robustify the system are required, e.g., [27] uses a combination of state and time delay observers and [28] implements a modified DMPC resilient against *denial of service* (DoS) attacks. Closely related, [29] designs a controller for CVP robust against faults causing loss of communication. The literature dealing with other attack types such as injection attacks seems more scarce, e.g., [30] deals with various malicious threats and proposes a robust consensus strategy relying on the availability of sufficient uncorrupted communication links. A larger body of work exists on control methods robust to additive faults such as [31], where an integrated fault tolerant control based on a *reduced order unknown input observer* (R-UIO) is presented, and others like [32], [33], and [34]. These approaches can in some cases be employed for robustness against cyber-attacks.

Received 10 May 2022; revised 19 June 2023 and 25 April 2024; accepted 31 October 2024. Date of publication 3 December 2024; date of current version 9 January 2025. This work was supported in part by Spanish Training Program for Academic Staff under Grant FPU17/02653 and in part by Spanish Ministry of Science and Innovation (MCIN)/Spanish State Research Agency (AEI)/10.13039/501100011033 Project C3PO-R2D2 under Grant PID2020-119476RB-I00. The Associate Editor for this article was P. Pisu. *(Corresponding author: Riccardo Maria Giorgio Ferrari.)*

Twan Keijzer is with the Royal Netherlands Aerospace Center, 1059 CM Amsterdam, The Netherlands (e-mail: tkeijzer1994@gmail.com).

Paula Chanfreut is with the Department of Mechanical Engineering, Eindhoven University of Technology, 5612 AZ Eindhoven, The Netherlands (e-mail: p.chanfreut.palacio@tue.nl).

José María Maestre is with the Department of Systems and Automation Engineering, University of Seville, 41004 Seville, Spain (e-mail: pepemaestre@us.es).

Riccardo Maria Giorgio Ferrari is with Delft Center for Systems and Control, Delft University of Technology, 2628 CD Delft, The Netherlands (e-mail: r.ferrari@tudelft.nl).

Digital Object Identifier 10.1109/TITS.2024.3503370

As can be seen, while previous works have explored the dynamic formation of coalitions and the adaptation to communication failures and minimum distance constraints, there remains challenges on the comprehensive handling of cyber-attack scenarios, particularly those affecting the integrity and reliability of communicated signals between vehicles in a platoon. Moreover, while some strategies have been proposed to detect and mitigate the impacts of cyber-attacks, there is a noticeable gap in developing a unified, robust control framework that can seamlessly transition between operational modes to properly response to detected cyber threats. In this regard, the consideration of scenario-based approaches to handle the uncertainties introduced by cyber-attacks and dynamically changing coalitions in a predictive control framework has not been thoroughly investigated. Such an approach has the potential to offer a more resilient and adaptive solution to maintaining platoon integrity and safety while minimizing the impact on performance and communication overhead.

In response to these identified gaps, this article introduces a novel scenario-based coalitional MPC framework that addresses the challenges of cyber-attack resilience and dynamic coalition management, ensuring safety and robust performance under a wide range of operational disruptions. In particular, we present a control algorithm for CVPs, which integrates a coalitional MPC for input optimization and an R-UIO based method for cyber-attack detection. The coalitions are constructed by disabling attacked communication links while making a trade-off between performance and communication costs on all other communication links. In this regard, our work follows works as [29], which note that platoons can also operate safely with less communication, albeit with degraded performance. The MPC problem is proven to be recursively feasible under the changing coalitions and we prove that the resulting CVP is free of crashes and conforms to a relaxed string stability requirement, for the considered scenarios include the uncertainty incurred by undetected attacks as well as by changing coalitions.

The rest of the article is organized as follows. Section II describes the problem including the vehicle dynamics, the concept of coalitions, the attack model, and a list of requirements for the proposed solution. Section III introduces the design of the R-UIO used for cyber-attack detection. Section IV presents the topology-switching rule and the formulation of the MPC problem solved by the coalitions. Section V provides the theoretical guarantees of safety and string stability for the proposed control scheme. Section VI presents numerical results on a CVP of 4 vehicles following a leader. Finally, Section VII provides conclusions and future research prospects.

Notation: $\xi(n|k)$ indicates the predicted value of variable ξ computed at time instant k for $k+n$; $\text{conv}(\mathcal{A})$ denotes the convex hull of set $\mathcal{A}$; $\text{sgn}(x)$ denotes the sign of x, with $\text{sgn}(0) = 0$. For a set $\mathcal{C}$, $Q_{\mathcal{C}} = [Q_i]_{i\in\mathcal{C}}$ denotes a block diagonal matrix with $|\mathcal{C}|$ blocks Q_i, where $|\mathcal{C}|$ is the cardinality of $\mathcal{C}$. Finally, $|\cdot|$ denotes the absolute value when referring to a scalar and the cardinality when referring to a set as used before.

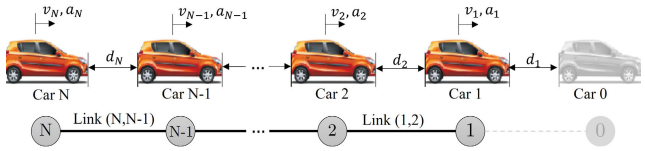

Fig. 1. Platoon of N vehicles following a lead vehicle. The structure of the communication network is modelled by graph $\mathcal{G} = \{\mathcal{N}, \mathcal{L}\}$.

II. Problem Formulation

We consider a CVP where vehicles can dynamically assemble into cooperative groups, hereafter referred to as *coalitions*. Local measurements are communicated between the members of each coalition to improve the control performance. However, this communication might be subject to cyber-attacks that can change the communicated signal, thus threatening the coalitions safety.

A. Vehicles Dynamics

Consider a CVP formed by a set $\mathcal{N} = \{1, \dots, N\}$ of locally controlled vehicles (see Figure 1). Each vehicle $i \in \mathcal{N}$ aims at keeping a reference distance

$$d_{r,i} \triangleq r + hv_i, \tag{1}$$

from its preceding vehicle $i-1$, where v_i is the velocity of vehicle i, h denotes the reference time headway and r is the reference distance at standstill.

Additionally, vehicle platoons require a form of string stability [35], for which we define two variants in time-domain as follows:

Definition 1 (Strict string stability): A vehicle platoon is strictly string stable if for every vehicle i and any given time instants k_0 and k_1, with $k_1 > k_0$, it holds that

$$\left|\frac{v_i(k_1) - v_i(k_0)}{v_{i-1}(k_1) - v_{i-1}(k_0)}\right| < 1.$$

Definition 2 (Relaxed string stability): A vehicle platoon is relaxed string stable if for every vehicle l and at any given time instants k_0 and k_1, with $k_1 > k_0$,

$$\exists\ i \text{ s.t. } \left|\frac{v_i(k_1) - v_i(k_0)}{v_l(k_1) - v_l(k_0)}\right| < 1 \text{ and } l < i.$$

Remark 1: Strict string stability assures that the impact of disturbances decreases between any two vehicles moving further away from the source of the disturbance. Relaxed string stability allows for bounded violations of strict string stability between any two vehicles, as long as after some number of vehicles the string stability property is regained. ◁

The continuous time dynamics of each vehicle $i \in \mathcal{N}$ are modelled as

$$\begin{cases} \dot{e}_{\text{d},i} &= -ha_i + \Delta v_i, \\ \dot{d}_i &= \Delta v_i, \\ \dot{v}_i &= a_i, \\ \dot{a}_i &= \dfrac{1}{\tau}(u_i - a_i), \\ \Delta\dot{v}_i &= a_{i-1} - a_i, \end{cases} \tag{2}$$

where $e_{\mathrm{d},i} \triangleq d_i - d_{r,i}$, d_i, a_i, Δv_i, and u_i are respectively the tracking error, distance, acceleration, relative velocity, and input of vehicle i. Furthermore, τ is the time constant of the vehicle drive-train. Note that the dynamics of each vehicle i is only affected by the preceding vehicle $i-1$ through a_{i-1}.

System (2) can be discretized using a timestep T and written in state space form as

$$\begin{cases} x_i(k+1) = A_{i,i}x_i(k) + B_{i,i}u_i(k) + A_{i,i-1}x_{i-1}(k), \\ y_i(k) \quad\quad = x_i(k), \end{cases} \tag{3}$$

where $x_i(k) = [e_{\mathrm{d},i}(k),\ d_i(k),\ v_i(k),\ a_i(k),\ \Delta v_i(k)]^\top \in \mathbb{R}^{n_x}$ and $u_i(k) \in [u_{\min}, u_{\max}] \in \mathbb{R}$ are the discrete-time state and input of vehicle i, and $y_i(k) \in \mathbb{R}^{n_y}$ represents its output vector. Note that the state x_i and input $u_{\min} \le u_i \le u_{\max}$ of each vehicle $i \in \mathcal{N}$ are bounded due to the physical limitations of the vehicles.

Assumption 1: Each vehicle $i \in \mathcal{N}$ measures its own velocity and acceleration. The distance and relative velocity with respect to the preceding vehicle is also measured, e.g., using LIDAR. ◁

B. Switching Communication Topologies

Following the *coalitional* control approach of [11] and [36], we assume that vehicles are interconnected by a set of wireless communication links that allow each vehicle i to exchange its local measurement y_i. These links are considered to be bidirectional, i.e., any pair of connected vehicles can both send and receive information to/from the other. Furthermore, we also consider multi-hop communication, i.e., vehicles connected by a path of enabled links can exchange data.

Communication links can be dynamically enabled and disabled, leading to different communication topologies. Any communication topology induces a partition of the set of vehicles into *disconnected* groups, known as coalitions. Considering this, let us introduce the following notation:

- Set $\mathcal{C} \subseteq \mathcal{N}$ denotes a coalition of vehicles, i.e., a group of consecutive vehicles that exchange data and coordinate their actions for their joint benefit.
- Λ denotes the topology of the communication network.
- Set $\mathcal{P}_\Lambda$ denotes the partition into coalitions associated with a certain communication topology Λ, i.e.,

$$\mathcal{P}_\Lambda = \{\mathcal{C}_1, \mathcal{C}_2, \ldots, \mathcal{C}_{|\mathcal{P}_\Lambda|}\}, \tag{4}$$

  where $\cup_{\mathcal{C}_i \in \mathcal{P}_\Lambda} \mathcal{C}_i = \mathcal{N}$ and $\mathcal{C}_i \cap \mathcal{C}_j = \emptyset$, for all $\mathcal{C}_i, \mathcal{C}_j \in \mathcal{P}_\Lambda$. Note that the number of coalitions $|\mathcal{P}_\Lambda|$ will range from 1, if all the vehicles cooperate, to $|\mathcal{N}|$, in case the cars operate in a decentralized fashion.

See Figure 2 for an illustration of these concepts.

C. Unreliable Data Exchange

Malicious vehicles, which share untrustworthy information within their coalition, may exist in the platoon. The performed attacks are considered additive without assumptions on their form, such that they represent a general class of *false data injection* (FDI) attacks, which includes bias injection, zero-dynamics, and replay attacks. Let Λ be the chosen

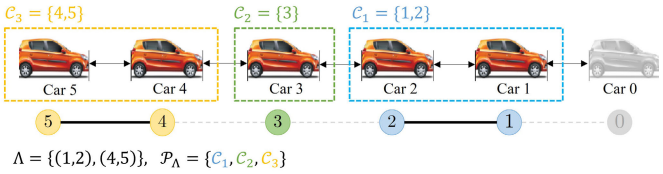

Fig. 2. Topology and resulting coalitions in a platoon with 5 vehicles following a leader vehicle which is not part of any coalition.

communication topology and $\mathcal{C} \in \mathcal{P}_\Lambda$ any of the resulting coalitions. Then, at each time instant k, vehicle $i \in \mathcal{C}$ receives the signals

$$y^i_j(k) = y_j(k) + a^i_{y_j}(k)$$

from each vehicle $j \in \mathcal{C} \setminus \{i\}$. Here $a^i_{y_j}(k)$ is the attack on the measurement vector sent from vehicle j to vehicle i. Note that for trustworthy vehicles it holds $a^i_{y_j}(k) = 0$.

Remark 2: Vehicles in different coalitions cannot attack each other because there is no inter-coalition communication. Therefore, working in a decentralized manner, i.e., when all coalitions are singletons, avoids the possibility of being attacked. A decentralized platoon, however, has lower performance due to the lack of coordination. ◁

D. Coalition Model

The model of a vehicle i as part of coalition $\mathcal{C}$ is[1]

$$\begin{cases} x_\mathcal{C}(k+1) &= A_\mathcal{C} x_\mathcal{C}(k) + B_\mathcal{C} u^i_\mathcal{C}(k) + w_\mathcal{C}(k)\,, \\ w_\mathcal{C}(k) &= A^w_\mathcal{C} x_{p\mathcal{C}}(k)\,, \\ x_{p\mathcal{C}}(k+1) &= A_{p\mathcal{C}} x_{p\mathcal{C}}(k) + B_{p\mathcal{C}} u_{p\mathcal{C}}(k)\,, \\ y^i_\mathcal{C}(k) &= C_\mathcal{C} x_\mathcal{C}(k) + C^i_\mathrm{a} a^i_{y\mathcal{C}}(k)\,, \end{cases} \tag{5}$$

where $x_\mathcal{C} = [x_j]_{j\in\mathcal{C}} \in \mathbb{R}^{n_x|\mathcal{C}|}$ is the aggregation of the states of all vehicles in $\mathcal{C}$, $u^i_\mathcal{C} = [u^i_j]_{j\in\mathcal{C}} \in \mathbb{R}^{|\mathcal{C}|}$ and $y^i_\mathcal{C} = [y^i_j]_{j\in\mathcal{C}} \in \mathbb{R}^{n_y|\mathcal{C}|}$ are respectively the coalitional input and output as known by vehicle $i \in \mathcal{C}$, $a^i_{y\mathcal{C}} = [a^i_{y_j}]_{j\in\mathcal{C}} \in \mathbb{R}^{n_y|\mathcal{C}|}$ are the aggregated attacks on vehicle i from all vehicles in the coalition. Furthermore $w_\mathcal{C} = [w_{\min(\mathcal{C})}(k), 0, \ldots, 0]^\top$ represents the coupling between the first vehicle in coalition $\mathcal{C}$, i.e., $\min(\mathcal{C})$, and vehicle $p_\mathcal{C}$ preceding the coalition. Notice that, given (2), variable $w_\mathcal{C}$ depends only on the acceleration of the preceding car, i.e., $a_{p\mathcal{C}}$. Here $p_\mathcal{C}$ is defined as $p_\mathcal{C} = \min(\mathcal{C}) - 1$. For example, in Figure 2, the predecessor of coalition 3, formed by cars 4 and 5, is vehicle 3, i.e. $p_{\mathcal{C}_3} = 3$. The matrices $A_\mathcal{C}$, $A_{p\mathcal{C}}$, $A^w_\mathcal{C}$, $B_\mathcal{C}$, $B_{p\mathcal{C}}$, and $C_\mathcal{C}$ are built according to Model (3), and C^i_a is a matrix that maps the attacks in $a^i_{y\mathcal{C}}$ into the corresponding components of $y^i_\mathcal{C}$.

As shown in Figure 2, the overall system can be seen as a sequence of cooperative substrings which respectively follow a vehicle whose actions are uncertain, yet bounded as $u_{\min} \le u_{p\mathcal{C}} \le u_{\max}$. Furthermore, due to the possibility of cyber-attacks, uncertainty also exists in the data communicated among vehicles. Using a combination of the cyber-attack detector and topology switching rule, which will be presented in Sections III and IV-A respectively, this effect of the attack

[1]For the sake of clarity, hereafter we use $\mathcal{C}$ to refer to a coalition in general, but note that there may be a number of different coalitions in the system simultaneously as indicated in (4).

TABLE I
VARIABLES DEFINITION

$e_{\mathrm{d},i}$	Tracking error associated with vehicle i	$a^i_{y_j}$	Attacks received by vehicle i on vehicle j's output vector
d_i	Distance from vehicle i to $i-1$	$\hat{\alpha}_i$	Estimated attack on the communication received by vehicle i
v_i	Speed of vehicle i	$\overline{x}_i$	Augmented state of vehicle i
a_i	Acceleration of vehicle i	e_i	Error between the estimated and real attack on the communication received by vehicle i
Δv_i	Relative speed associated with vehicle i ($\Delta v_i = v_{i-1} - v_i$)	$\mathcal{S}, s$	Set of scenarios used in the the MPC problem $\mathcal{S}$ with instances s
Δa_i	Relative acceleration associated with vehicle i ($\Delta a_i = a_{i-1} - a_i$)	$\hat{\mathbf{u}}_{p_\mathcal{C},s}$	Input sequence of the vehicle preceding coalition $\mathcal{C}$ according to scenario s
x_i, u_i, y_i	State, input and output of vehicle i	$\hat{a}_{y_\mathcal{C},s}$	Attack on the communication between vehicles in coalition $\mathcal{C}$ according to scenario s
y^i_j	Output vector of vehicle j received by vehicle i	T	Discrete time step
$\mathcal{C}_i$	Coalition i	T_α	Threshold for detection of attacks
$p_\mathcal{C}$	Vehicle preceding coalition $\mathcal{C}$	T_d	Threshold for starting communication based on the distance tracking error between vehicles
$\mathbf{u}_\mathcal{C}$	Input sequence of coalition $\mathcal{C}$ optimized agents $i \in \mathcal{C}$ over a prediction horizon	T_v	Threshold for starting communication based on the relative velocity between vehicles

Note: When subscript i or j refers to a single vehicle, subscript $\mathcal{C}$ indicates the same variable stacked for all vehicles in $\mathcal{C}$.

can be bounded in a convex set $\mathcal{A}_\mathcal{C}$, which will be defined later.

Remark 3: Since the cyber-attack $a^i_{y_\mathcal{C}}(k)$ affects the measurement vector $y^i_\mathcal{C}$, it can affect the computation of $u^i_\mathcal{C}(k)$ through the controller which will be introduced in Section IV, and thus affect the vehicle behaviour. ◁

E. Requirements and Proposed Solution

As shown in Figure 3, the proposed control scheme comprises three main components:

1) An R-UIO, which must estimate and detect cyber-attacks on the communication within each coalition.
2) A topology-switching law, pursuing the following goals:
   a) Disable communication links when they do not provide significant performance improvements.
   b) Disable communication links when cyber-attacks are detected by the R-UIO.
   c) Guarantee that the platoon is sufficiently connected such that relaxed string stability holds at a CVP level.
3) The coalitional MPC, which is designed to:
   a) Provide optimal reference tracking control robust against undetected attacks and uncertain actions of preceding vehicle $p_\mathcal{C}$.
   b) Avoid crashes for all vehicles, even when communication links are attacked, i.e., $d_i(k) > 0,\ \forall i \in \mathcal{N},\ k \geq 0$.
   c) Guarantee that there always exists a feasible input for each vehicle such that, in healthy conditions, strict string stability holds within each coalition.

Remark 4: The proposed control scheme is designed to guarantee safety against the FDI attack described in Section II-C. It can however also guarantee safety from denial of service (DoS) attacks. In this case, no communication is received such that attack detection is trivial. Furthermore, as mitigation is achieved by disabling the affected communication links, this step is redundant for DoS attacks. The MPC can then be used without change to achieve safety from DoS attacks.

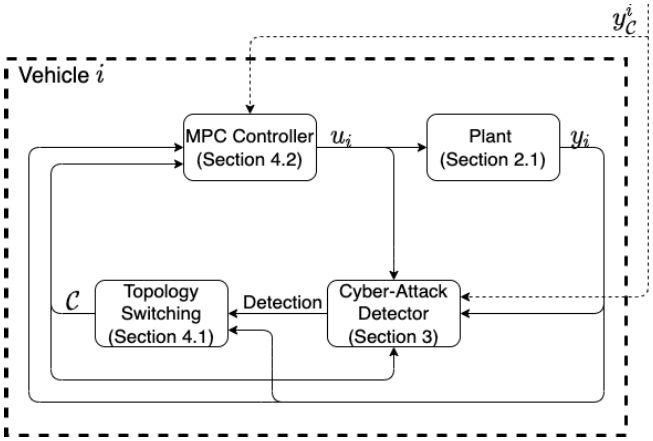

Fig. 3. Block diagram of the control solution used in each vehicle. The dotted arrows indicate signals communicated from other vehicles in the coalition.

III. CYBER-ATTACK DETECTOR DESIGN

Each vehicle uses the cyber-attack detection method based on the R-UIO from [31]. A discretized version of this R-UIO is presented in this section along with a detection threshold and guarantees on its performance. As can be seen in Figure 3, when an attack is detected, a signal is sent to the topology switching module, which will then disable the corresponding communication link.

The system in Equation (5) can be augmented by aggregating the state and the attack as

$$\begin{cases} \bar{x}_i(k+1) & = \bar{A}_\mathcal{C}\bar{x}_i(k) + \bar{B}_\mathcal{C}u^i_\mathcal{C}(k) + \bar{d}_i(k)\,, \\ y^i_\mathcal{C}(k) & = \bar{C}_i\bar{x}_i(k)\,, \end{cases} \tag{6}$$

where

$$\bar{x}_i(k) = \begin{bmatrix} x_\mathcal{C}(k) \\ a^i_{y_\mathcal{C}}(k) \end{bmatrix},\ \bar{d}_i(k) = \begin{bmatrix} w_\mathcal{C}(k) \\ a^i_{y_\mathcal{C}}(k+1) \end{bmatrix},\ \bar{B}_\mathcal{C} = \begin{bmatrix} B_\mathcal{C} \\ 0 \end{bmatrix},$$
$$\bar{A}_\mathcal{C} = \begin{bmatrix} A_\mathcal{C} & 0 \\ 0 & 0 \end{bmatrix},\ \bar{C}_i = \begin{bmatrix} I & C^i_\mathrm{a} \end{bmatrix}.$$

Here, we define the sets $\mathcal{D}_i = \{\bar{d}_i | w_\mathcal{C} \in \mathcal{W}_\mathcal{C}, a^i_{y_\mathcal{C}} \in \mathcal{A}_\mathcal{C}\}$ and $\mathcal{D}^0_i = \{\bar{d}_i | w_\mathcal{C} \in \mathcal{W}_\mathcal{C}, a^i_{y_\mathcal{C}} = 0\}$ as the possible disturbances $\bar{d}_i$

in, respectively, attacked and healthy conditions. For a system of the form (6), the R-UIO to estimate the attack signals $a^i_{y_\mathcal{C}}(k) = L\bar{x}_i(k)$, with $L = [0\ I]$, can be written as

$$\begin{cases} \hat{x}_{\alpha_i}(k+1) = M_i\hat{x}_{\alpha_i}(k) + G_i u^i_{\mathcal{C}}(k) + R_i y^i_{\mathcal{C}}(k)\,, \\ \hat{\alpha}_i(k) = \hat{x}_{\alpha_i}(k) + H_i y^i_{\mathcal{C}}(k)\,, \end{cases} \tag{7}$$

where $\hat{\alpha}_i$ estimates $a^i_{y_\mathcal{C}}$ and

$$\begin{aligned} ||M_i|| &< 1\,, \\ M_iT_i + R_i\bar{C}_i - T_i\bar{A}_\mathcal{C} &= 0\,, \\ T_i + H_i\bar{C}_i - L &= 0\,, \\ G_i - T_i\bar{B}_\mathcal{C} &= 0\,. \end{aligned} \tag{8}$$

Here, the matrices in Equation (7) are designed based on System (6) to satisfy the constraints in Equation (8) as in [31]. To achieve stability of the discrete-time observer, only the condition on M_i is changed w.r.t. [31]. Furthermore, T_i appears in the full derivations of the R-UIO as a transformation matrix, and can be freely designed such that the conditions in (8) hold. Then, the observer error dynamics are reduced to

$$e_i(k+1) = M_i e_i(k) + (H_i\bar{C}_i - L)\bar{d}_i(k), \tag{9}$$

where $e_i = \hat{\alpha}_i - L\bar{x}_i(k)$.

Let us now define a threshold T_α for attack detection as

$$\begin{aligned} T_\alpha(k) &\triangleq ||(I - M_i)^{-1}||D^0_i, \\ D^0_i &= \max_{\bar{d}_i \in \mathcal{D}^0_i} ||(H_i\bar{C}_i - L)\bar{d}_i||, \end{aligned}$$

such that detection is triggered when

$$||\hat{\alpha}_i(k)|| > T_\alpha(k)\,. \tag{10}$$

Lemma 1 (Robustness): The threshold T_α is robust to uncertainties and does not lead to false detections.

Proof: First, under healthy conditions Equation (9) can be simplified as

$$\hat{\alpha}_i(k+1) = M_i\hat{\alpha}_i(k) + (H_i\bar{C}_i - L)\bar{d}_i(k)\,.$$

Now, if we initialize $\hat{\alpha}_i(0) = 0$, $\hat{\alpha}_i(k)$ can be written as

$$\hat{\alpha}_i(k) = \sum_{j=0}^{k-1} M_i^j(H_i\bar{C}_i - L)\bar{d}_i(j) \tag{11}$$

so that

$$||\hat{\alpha}_i(k)|| \le \sum_{j=0}^{k-1} ||M_i^j||D^0_i \le ||(I - M_i)^{-1}||D^0_i = T_\alpha(k)\,,$$

which concludes the proof. □

To analyze detectability under attacks, let us define

$$D_i \triangleq \max_{\bar{d}_i \in \mathcal{D}_i} ||(H_i\bar{C}_i - L)\bar{d}_i||\,. \tag{12}$$

Theorem 1 (Detectability): A sufficient condition for attack detection by the R-UIO is $a^i_{y_\mathcal{C}}(k) \notin \mathcal{A}_\mathcal{C}$, with

$$\mathcal{A}_\mathcal{C} \triangleq \{a^i_{y_\mathcal{C}} : ||a^i_{y_\mathcal{C}}|| \le ||(I - M_i)^{-1}||(D_i + D^0_i)\}$$

being the set of attacks not guaranteed to be detected.

Proof: Using Equation (9), the disturbance bound in Equation (12), and assuming the UIO is initialized in healthy conditions, i.e., $e_i(0) = 0$, we can derive

$$||e_i(k)|| \le ||(I - M_i)^{-1}||D_i\,.$$

This implies

$$||\hat{\alpha}_i(k)|| \ge ||a^i_{y_\mathcal{C}}(k)|| - ||(I - M_i)^{-1}||D_i.$$

Using this in combination with the detection condition from Equation (10), detection is guaranteed if

$$||a^i_{y_\mathcal{C}}(k)|| - ||(I - M_i)^{-1}||D_i > T_\alpha(k),$$

which, by definition of T_α, proves the theorem. □

Corollary 1: The set of attacks for which detection is not guaranteed can be over-bounded by a convex polytope $\bar{\mathcal{A}}_\mathcal{C} \triangleq \text{conv}(\{\alpha_0, \ldots, \alpha_\eta\}) \supset \mathcal{A}_\mathcal{C}$, were all $\alpha_\bullet \in \mathbb{R}^{n_y|\mathcal{C}|}$.

Remark 5: Corollary 1 implies the existence of *stealthy attacks* [37], that is attacks that cannot be detected using the proposed detector. Still, the Corollary provides a characterization via the set $\bar{\mathcal{A}}_\mathcal{C}$, which will be used in Section IV-B to define a scenario-based MPC problem that can guarantee safety even in the presence of a stealthy attack.

IV. Topology Switching Controller

In this section, the coalitional MPC controller is presented. First, we describe the topology switching rule, which sets the number and composition of the coalitions, and, secondly, we formulate the MPC problem to be solved by each vehicle.

A. Topology Switching Rule

The communication topology and the coalitions are dynamically updated to attain a trade-off between optimal performance and coordination efforts while mitigating the effect of attacks. In particular, they are selected according to the tracking error, relative velocities, and attacks estimation as described in Algorithm 1. Note that, if attacks are not detected, communication and coordination between vehicles is enabled if the tracking errors or relative velocities exceed thresholds T_d and T_v (see lines 4 and 5 of Algorithm 1). Otherwise, sparser communication topologies are imposed by placing the vehicles in different coalitions (see line 7 of Algorithm 1). The thresholds T_v and T_d affect the communication topology in nominal conditions. Setting higher thresholds will cause less communication and computation costs, but also a lower nominal control performance. In the same way, lower thresholds will cause more communication and computation costs and higher control performance.

However, when the communication is subject to cyber-attacks, the main goal shifts to preserving safety. Communication links on which a cyber-attack is performed are dangerous to maintain and, therefore, they are disabled upon notifications from the detector of Section III. By doing so, Theorem 1 implies that only attacks $a^i_{y_\mathcal{C}} \in \mathcal{A}_\mathcal{C}$ can affect the platoon. Note that, while Algorithm 1 could be implemented by a *central* coordinator, centralized computations are not needed. Indeed, each car decides when to enable/disable the link with

the vehicle in front or in the back according to the mentioned criteria.

Remark 6: To understand this point, let us note that Algorithm 1 is a constructive algorithm. In line 1 all coalitions are initialized, regardless of their previous state, to only a coalition including only vehicle 1. Then, for each vehicle, a test coalition $\mathcal{C}_\star$ is considered for the algorithm purpose only, which includes the current vehicle and the previous one (line 3). Both vehicles in $\mathcal{C}_\star$ can then compute their estimate $\hat{\alpha}$ and compare it to the threshold. Only if both estimates are lower than the threshold, which means that both vehicles trust each other, and the test on the relative velocity and position error is passed, then (line 6) the current vehicle is added to a coalition with the previous one(s). As an example, let us consider three vehicles, of which the second one is malicious. At initialization we would have $\mathcal{C}_1 = \{1\}$ only. When $i = 2$ the test coalition is $\mathcal{C}_\star = \{1, 2\}$ and vehicle 1 will possibly detect vehicle 2 as malicious, thus not adding it to coalition $\mathcal{C}_1$. Vehicle 2 will thus end up being alone in $\mathcal{C}_2 = \{2\}$ (line 7). At next step $i = 3$, $\mathcal{C}_\star = \{2, 3\}$ and vehicle 3 will possibly detect 2 as malicious. This will cause vehicle 3 to end up, alone, in $\mathcal{C}_3 = \{3\}$ (line 7). While in this case no collaboration occurs, the important goal reached is that the malicious vehicle 2 is not added to any coalition, thus preserving safety.

B. MPC Problem

Vehicles within each coalition are affected by uncertainty through the actions of the vehicle preceding the coalition and potential undetected cyber-attacks on the communicated signals. In particular, attacks may not be detected if they belong in the set $\mathcal{A}_\mathcal{C}$ defined in Theorem 1. Safety in this case is guranteed using a scenario-based MPC that uses the convex polytope $\bar{\mathcal{A}}_\mathcal{C}$ introduced in Corollary 1. A detected attack, on the other hand, will lead to that communication link being immediately disabled.

Scenario-based approaches consider a set of realizations of the uncertainties affecting the system. The MPC problem is formulated so that the implemented inputs satisfy the system constraints in these scenarios, while optimizing an objective function that typically weights the performance costs in all of these situations. Although the scenario-based approach usually provides stochastic guarantees on constraints satisfaction, here we consider the extreme realizations of the vehicles' behaviour and undetected attacks, such that safety guarantees in all cases can be obtained.

Algorithm 1 Topology Switching Rule

1: **Initialize:** $\mathcal{C}_1 = \{1\}$ and $j = 1$ and $\mathcal{C}_2 = \ldots \mathcal{C}_N =$
2: **for all** vehicles $i = 2 \ldots N$ **do**
3: Set $\mathcal{C}_\star = \{i - 1, i\}$.
4: **if** $(|\Delta v_i| > T_\mathrm{v}$ **or** $|e_{\mathrm{d},i}| > T_\mathrm{d})$ **and** $\|\hat{\alpha}_i(k)\| \leq T_\alpha(k)$ **and** $\|\hat{\alpha}_{i-1}(k)\| \leq T_\alpha(k)$
5: Set $\mathcal{C}_j = \{\mathcal{C}_j, i\}$.
6: **else**
7: Set $j = j + 1$ and $\mathcal{C}_j = \{i\}$.
8: **end if**
9: **end for**

1) Uncertainty Scenarios: At each time instant k, each vehicle $i \in \mathcal{C}$ considers a set of S realizations on the unknown neighboring variables. In particular, each scenario $s \in \mathcal{S} = \{1, \ldots, S\}$ defines possible undetected attacks on the measurement vector, $\hat{a}^i_{y_\mathcal{C},s}$ and a possible trajectory of the coalitions' predecessor input, i.e.,

$$\hat{\mathbf{u}}_{p_\mathcal{C},s} = [\hat{u}_{p_\mathcal{C},s}(k|k), \ldots, \hat{u}_{p_\mathcal{C},s}(k + N_\mathrm{p} - 1|k)]\,.$$

Here N_p is the length of the prediction horizon. As used above, in what follows, let us use subscript s to indicate the scenarios, e.g., $x_{\mathcal{C},s}(n|k)$ will denote the prediction made at time instant k for the state of coalition $\mathcal{C}$ in scenario s at time instant n.

We assume that the set of scenarios $\mathcal{S}$ can be divided into three different categories. First, we define $\mathcal{S}_\mathrm{e} \subset \mathcal{S}$ as the subset of *extreme* scenarios, which imply that the predecessor input and the undetected cyber-attacks take their extreme values, i.e.,

$$\mathcal{S}_\mathrm{e} = \{s \in \mathcal{S} \quad \hat{u}_{p_\mathcal{C},s}(n|k) \in \begin{cases} \{u_\mathrm{min}, u_\mathrm{max}\} & \text{if } v_{p_\mathcal{C}}(n|k) \in [0, v_\mathrm{max}], \\ 0 & \text{otherwise,} \end{cases}$$
$$\hat{a}^i_{y_\mathcal{C},s}(k|k) \in \{a_0, \ldots, a_\eta\},$$
$$n = k, \ldots, k + N_\mathrm{p} - 1\}.$$

These scenarios are used to guarantee safety of the CVP. Secondly, we consider a set of *healthy extreme* scenarios $\mathcal{S}_0 \subset \mathcal{S}$, which involve only extreme inputs, while the cyber-attack vector is zero. That is,

$$\mathcal{S}_0 = \{s \in \mathcal{S} \quad \hat{u}_{p_\mathcal{C},s}(n|k) \in \begin{cases} \{u_\mathrm{min}, u_\mathrm{max}\} & \text{if } v_{p_\mathcal{C}}(n|k) \in [0, v_\mathrm{max}], \\ 0 & \text{otherwise,} \end{cases}$$
$$\hat{a}^i_{y_\mathcal{C},s}(k|k) = 0,$$
$$n = k, \ldots, k + N_\mathrm{p} - 1\big\}\,.$$

These healthy extreme scenarios are used for string stability in healthy conditions. Lastly, other scenarios, denoted as *design* scenarios $\mathcal{S}_\mathrm{d}$, can be chosen freely to include other hypotheses on the realizations of the uncertainty, i.e.,

$$\mathcal{S}_\mathrm{d} = \{s \in \mathcal{S} \quad u_\mathrm{min} \leq \hat{u}_{p_\mathcal{C},s}(n|k) \leq u_\mathrm{max},$$
$$\hat{a}^i_{y_\mathcal{C},s}(k|k) \in \bar{\mathcal{A}}_\mathcal{C},$$
$$n = k, \ldots, k + N_\mathrm{p} - 1\big\}\,.$$

Therefore, the user can add any finite number of *design* scenarios at the expense of an increase in computational burden. Finally, note that $\mathcal{S} = \mathcal{S}_\mathrm{e} \cup \mathcal{S}_0 \cup \mathcal{S}_\mathrm{d}$.

2) Ideal MPC Problem: The ideal MPC problem considers the safety and string-stability conditions that we want to satisfy, but, as will be shown, it cannot be used directly for real-time control. Modifications to make this possible will lead to the *practical* MPC problem of subsection IV-B.3.

The ideal MPC problem can be formulated as follows:

$$\min_{\mathbf{u}^i_\mathcal{C}} \; J_\mathcal{C}(y^i_\mathcal{C}(k), \mathbf{u}^i_\mathcal{C}) \tag{13}$$

subject to:

Prediction model

$$x_{\mathcal{C},s}(k|k) = y^i_\mathcal{C}(k) - C^i_a \hat{a}^i_{y_\mathcal{C},s}(k|k), \tag{14a}$$

$$x_{\mathcal{C},s}(n+1|k) = A_{\mathcal{C}} x_{\mathcal{C},s}(n|k) + B_{\mathcal{C}} u^i_{\mathcal{C}}(n|k) + w_{\mathcal{C},s}(n|k), \quad (14b)$$

$$w_{\mathcal{C},s}(n|k) = A^w_{\mathcal{C}} x_{p_{\mathcal{C}},s}(n|k), \quad (14c)$$

$$x_{p_{\mathcal{C}},s}(k|k) = x_{p_{\mathcal{C}},s}(k|k-1), \quad (14d)$$

$$x_{p_{\mathcal{C}},s}(n+1|k) = A_{p_{\mathcal{C}}} x_{p_{\mathcal{C}},s}(n|k) + B_{p_{\mathcal{C}}} \hat{u}_{p_{\mathcal{C}},s}(n|k), \quad (14e)$$

$$u^i_{\mathcal{C}}(n|k) \in [u_{\min} \quad u_{\max}]^{|\mathcal{C}|}, \quad \forall s \in \mathcal{S}, \quad (14f)$$

Safety

$$d_{i,s}(n|k) \geq 0 \ \forall s \in \mathcal{S}_{\mathrm{e}}, \quad (15)$$

String stability

$$\mathrm{sgn}(\Delta v_{i,s}(n|k)) = \mathrm{sgn}(\mathrm{d}v_{i,s}(k|k)) \ \forall s \in \mathcal{S}_0, \quad (16)$$

$$\forall n = k, \ldots, k + N_{\mathrm{p}} - 1,$$

where cost function $J_{\mathcal{C}}(y^i_{\mathcal{C}}(k), \mathbf{u}^i_{\mathcal{C}})$ is of the form

$$J_{\mathcal{C}}(y^i_{\mathcal{C}}(k), \mathbf{u}^i_{\mathcal{C}}) = \sum_{n=k}^{k+N_{\mathrm{p}}-1} \left(\sum_{s \in \mathcal{S}_{\mathrm{d}}} p_s x_{\mathcal{C},s}(n+1|k)^{\mathrm{T}} Q_{\mathcal{C}} x_{\mathcal{C},s}(n+1|k) + \Delta u^i_{\mathcal{C}}(n|k)^{\mathrm{T}} R_{\mathcal{C}} \Delta u^i_{\mathcal{C}}(n|k) \right).$$

Here $Q_{\mathcal{C}} = [Q_i]_{i \in \mathcal{C}}$ and $R_{\mathcal{C}} = [R_i]_{i \in \mathcal{C}}$ are positive definite weighting matrices defined as the block-diagonal aggregation of Q_i and R_i, respectively, and $p_s > 0$ represents the probability assigned to scenario $s \in \mathcal{S}_d$. Furthermore, $\mathbf{u}^i_{\mathcal{C}}$ is the sequence vector $\mathbf{u}^i_{\mathcal{C}} = [{u^i_{\mathcal{C}}}^{\top}(k|k), \ldots, {u^i_{\mathcal{C}}}^{\top}(k+N_{\mathrm{p}}-1|k)]^{\top}$, and $\Delta u^i_{\mathcal{C}}(n|k)$ is defined as $\Delta u^i_{\mathcal{C}}(n|k) = u^i_{\mathcal{C}}(n|k) - u^i_{\mathcal{C}}(n-1|k)$.[2] Finally, $\mathrm{d}v_{i,s}(k|k) = v_{i,s}(k+N_{\mathrm{p}}|k) - v_{i,s}(k|k)$ denotes the predicted change of velocity of vehicle i over the prediction horizon. Note that unlike a min-max approach, here the deterministic worst case scenarios $\mathcal{S}_{\mathrm{e}}$ and $\mathcal{S}_0$ are used to guarantee safety in terms of constraint satisfaction, but the minimization is not performed based on the worst case scenario.

In this ideal MPC problem, (14a)–(14f) predict the coalition behaviour over the prediction horizon for a given $s \in \mathcal{S}$. If scenario s occurs, then the behaviour predicted by (14a)–(14f) will be accurate and the attack will have no effect because we are essentially *subtracting* it in (14a). Note that, differently than a fault, the attacks considered in this article do not directly affect the predicted dynamics but only the initial condition for the prediction. That is, an attacker can modify the information it communicates to the other vehicles in its coalition, and thus the data used by the latter to determine the current coalition state (recall Section II-C). Therefore, if agent i is attacked, it will translate into dealing with a misleading $y^i_{\mathcal{C}}(k)$ in (14a).

Furthermore, the following lemma can be proved for the ideal MPC problem.

Lemma 2: For a fixed communication topology, if constraints (15) and (16) hold, we have:

- No crashes in the platoon, even when the system is under attack.
- Strict string stability guarantees within each coalition for the healthy system.

Proof: If constraint (15) holds, then for all extreme scenarios it holds $d_{i,s} > 0$. This implies there are no crashes for all possible uncertainties, including those from undetected cyber-attacks.

Constraint (16) guarantees strict string stability according to Definition 1 within each coalition for the healthy system. This can easily be derived by noting that, starting from $\Delta v_i(k) = 0$, Constraint (16) implies that if $\mathrm{d}v_{i,s}(k|k) > 0$, then $v_{i-1,s}(n|k) > v_{i,s}(n|k)$ and thus $\mathrm{d}v_{i-1,s}(k|k) > \mathrm{d}v_{i,s}(k|k)$. Conversely if $\mathrm{d}v_{i,s}(k|k) < 0$ then $\mathrm{d}v_{i-1,s}(k|k) < \mathrm{d}v_{i,s}(k|k)$. □

Unfortunately, Problem (13) cannot be readily implemented to find the vehicles' inputs in real-time. Firstly, the string stability Constraint (16) is non-linear, complicating the solution of the ideal MPC problem in real-time. Secondly, both the safety Constraint (15) and the string stability Constraint (16) are not recursively feasible for all scenarios. For these reasons, we propose a *modification* of Problem (13), which, at the expense of a certain loss of optimality, results in a recursively feasible quadratic optimization with linear constraints.

3) Practical MPC Problem: To obtain recursive feasibility through quadratic optimization with linear constraints, we need to reformulate the safety constraint and string stability constraint.

Safety constraints: To make the ideal safety Constraint (15) recursively feasible, it needs to be extended so that a feasible solution exists in all scenarios, including emergency braking of the car preceding the platoon and any undetected attack on the communication. To achieve the required robustness to uncertainty, the distance between vehicles is bounded based on the relative velocity and acceleration between vehicles, so that the preceding vehicle is not approached too fast. The exact relation of the practical safety constraint is

$$d_{j,s}(m|k) \geq 0, \quad (17a)$$

$$d_{j,s}(m|k) \geq -\Delta v_{j,s}(m|k)\delta(m|k), \quad (17b)$$

$$d_{j,s}(m|k) \geq -\left(\Delta v_{j,s}(m|k) + \tau \Delta a_{j,s}(m|k)\right)\delta(m|k), \quad (17c)$$

for scenarios $s \in \mathcal{S}_{\mathrm{e}}$, $m = k+1$, and for all $j \in \mathcal{C}$. Here, $\Delta a_{j,s} = a_{j-1,s} - a_{j,s}$ and

$$\delta(n|k) = \gamma(k|k) - (n-k)T,$$

with $\gamma(k|k)$ the time to standstill as defined below and $n \geq k$. Note that depending on the sign of the relative velocities and accelerations, there will be always one of the three conditions in (17) more conservative than the others. The latter will be used in Section V to guarantee safety and recursive feasibility even if the coalitions break into smaller ones.

Definition 3: $\gamma(k|k)$ is an upper bound on the time to standstill of vehicle j, when $u_j(\kappa) = u_{\min}$ $\forall \kappa \geq k$.

Remark 7: $\gamma(k|k)$ can be implicitly calculated through Model (3) given initial conditions $v_j(k)$ and $a_j(k)$. ◁

The constraints in (17) are all based on the idea that $d_j(n) > d_j(k) + \min_{k \leq \kappa < n}(\Delta v_j(\kappa))\gamma(k|k) > 0$, i.e. the change in distance can be bounded by a product of bounds on the relative velocity and the time to standstill. This relation is expanded

[2]Note that to obtain $\Delta u^i_{\mathcal{C}}(k|k)$, it is considered that $u^i_{\mathcal{C}}(k-1|k) = u^i_{\mathcal{C}}(k-1|k-1)$.

for three situations. In boundary case $d_{j,s}(n|k) = 0$, when Constraint (17a) is active, the relative velocity can only be positive. In the other cases sufficient distance must be held to guarantee recursive feasibility. Constraint (17b) is active only if the relative velocity is negative and the relative acceleration is positive, and Constraint (17c) is active only if both the relative velocity and acceleration are negative.

The full proof of recursive feasibility of this constraint is deferred to that of Theorem 2 in the next section.

String stability: The ideal string stability Constraint (16) is both non-linear and there are no guarantees that it can be recursively satisfied. Therefore a major reformulation of the constraint is required for the practical MPC problem. First, let us define the positive and negative components of $\mathrm{d}v_{j,s}$ as

$$v_{j,s}(k+N_\mathrm{p}|k) - v_{j,s}(k|k) = \mathrm{d}v_{j,s}^{\mathrm{pos}} + \mathrm{d}v_{j,s}^{\mathrm{neg}}, \tag{18a}$$

$$\mathrm{d}v_{j,s}^{\mathrm{pos}} \geq 0, \;\; \mathrm{d}v_{j,s}^{\mathrm{neg}} \leq 0\,. \tag{18b}$$

Furthermore, to assure that $\mathrm{d}v_{j,s}^{\mathrm{pos}}$ and $\mathrm{d}v_{j,s}^{\mathrm{neg}}$ will not unnecessarily cancel each other, we also add a term $\beta_1(\mathrm{d}v_{j,s}^{\mathrm{pos}} - \mathrm{d}v_{j,s}^{\mathrm{neg}})$ to the cost function. With this, we could obtain a linear constraint equivalent to Constraint (16) as

$$\gamma \mathrm{d}v_{j,s}^{\mathrm{neg}}(k|k) \leq \Delta v_{j,s}(n|k) \leq \gamma \mathrm{d}v_{j,s}^{\mathrm{pos}}(k|k).$$

for all $n = k \ldots k+N_\mathrm{p}-1$, where γ is a sufficiently large constant. This constraint is however still not recursively feasible, as implicitly it does not allow for sign changes of $\Delta v_{j,s}$ and $\mathrm{d}v_{j,s}$. This is because $\mathrm{d}v_{j,s}(k|k)$ is required to have the same sign as all relative velocities over the prediction horizon $\Delta v_{j,s}(n|k)$, including the current relative velocity $\Delta v_{j,s}(k|k)$. The current relative velocity is fixed, and therefore, also the sign of $\mathrm{d}v_{j,s}(k|k)$ and $\Delta v_{j,s}(n|k)$ $\forall n = k \ldots k+N_\mathrm{p}-1$ cannot be changed. This repeats for each next prediction horizon such that the sign of $\Delta v_{j,s}$ and $\mathrm{d}v_{j,s}$ can never change.

To this end, the constraint is changed to

$$N_\mathrm{p} T \gamma \, \mathrm{d}v_{j,s}^{\mathrm{neg}} \leq d_{j,s}(k+N_\mathrm{p}|k) - d_{j,s}(k|k) \leq N_\mathrm{p} T \gamma \, \mathrm{d}v_{j,s}^{\mathrm{pos}}\,.$$

Here, $d_{j,s}(k+N_\mathrm{p}|k) - d_{j,s}(k|k) = \sum_{n=k}^{k+N_\mathrm{p}-1} \Delta v_{j,s}(n|k) T$ so that this constraint is equivalent to the preceding one if the sign of $\Delta v_{j,s}$ is constant over the prediction horizon. During normal operation, the sign of $\Delta v_{j,s}$ is constant except when the platoon transitions between accelerating and decelerating or vice versa. A proof that string stability can be achieved, even when such a transition occurs, is presented in Theorem 3.

Lastly, a sensible value for the design constant γ is chosen. From the distance reference defined in Equation (1), it can be seen that with any change in velocity $\mathrm{d}v_j$, the reference distance changes with $h\mathrm{d}v_j$. Therefore, we set γ such that $N_\mathrm{p} T \gamma = h$ and the distance between vehicles never changes more than required to track the reference. This means the controller will not overshoot the distance reference, and thus the relative velocity will not change sign during a continuous acceleration/deceleration maneuver. This gives the final constraint

$$h\,\mathrm{d}v_{j,s}^{\mathrm{neg}} - \epsilon_s \leq d_{j,s}(k+N_\mathrm{p}|k) - d_{j,s}(k|k) \leq h\,\mathrm{d}v_{j,s}^{\mathrm{pos}} + \epsilon_s \tag{18c}$$

where ϵ_s is a slack variable. This constraint is applied for scenarios $s \in \mathcal{S}_0$ and for vehicles $j \in \mathcal{C} \setminus \min(\mathcal{C})$, i.e. (18) is not applied to the first vehicle of every coalition. Therefore, it only guarantees strict string stability within each coalition and not between coalitions. It will be shown in Section V that, together with the proposed topology switching rule, it is possible to guarantee relaxed string stability over the entire platoon.

Using the above, at each time instant k, each vehicle $i \in \mathcal{C}$ solves an MPC optimization problem formulated as follows:

$$\begin{aligned}
\min_{\mathbf{u}_\mathcal{C}^i, \mathrm{d}v_{i,\mathrm{pos}}, \mathrm{d}v_{i,\mathrm{neg}}, \epsilon_s} \quad & J_\mathcal{C}(y_\mathcal{C}^i(k), \mathbf{u}_\mathcal{C}^i) \\
& + \sum_s (\beta_1(\mathrm{d}v_{i,s}^{\mathrm{pos}} - \mathrm{d}v_{i,s}^{\mathrm{neg}}) + \beta_2 \epsilon_s) \\
\text{s.t. } & (14), \;\; \forall s \in \mathcal{S}, \\
& (17), \;\; \forall j \in \mathcal{C}, \forall s \in \mathcal{S}_\mathrm{e}, \\
& (18), \;\; \forall j \in \mathcal{C} \setminus \min(\mathcal{C}), \forall s \in \mathcal{S}_0, \\
& \forall n = k, \ldots, k+N_\mathrm{p}-1,
\end{aligned} \tag{19}$$

where β_1 and β_2 are weighting factors for slack variables used in Constraint (18).

Remark 8: Note that Problem (19) can be solved locally by each vehicle $i \in \mathcal{C}$ once all vectors y_j^i, for $j \in \mathcal{C}/\{i\}$, are received. ◁

Remark 9: Increasing the length of the prediction horizon and/or the size of coalition $\mathcal{C}$ would consequently increase the number of optimization variables in Problem (19), and therefore the computational burden for all $i \in \mathcal{C}$. In particular, while the number of decision variables of the quadratic program grows linearly with the number of vehicles and the prediction horizon considered, the number of constraints depends on the number of scenarios considered, which has worst-case exponential growth. However, many of these constraints are redundant and there are universal bounds for scenario sampling that provide practical certainty regarding constraint satisfaction and depend linearly on the number of decision variables [38]. Therefore, the time required to solve the optimization problem could be expected to grow at a manageable rate with the coalition size and the prediction horizon in a practical setting. Likewise, note that it is straightforward to introduce a limit on the coalitions sizes in the topology switching rule (see Section IV-A). Imposing such a limit may result in diminished overall performance, yet it brings reduced computational costs, thereby facilitating the application of the proposed approach for real-time control. Finally, notice that lighter options than having each $i \in \mathcal{C}$ solving a coalition-wide could have been used to attain coordinated decisions, e.g., by using DMPC. However, these methods are typically iterative and introduce new communication steps which represent additional sources of vulnerability. To avoid this issue, we preferred to have a relatively larger problem size that could be solved only once.

Remark 10: Being a safety critical application, the proposed MPC problem is conservative. Although the cost in the MPC problem is not minimized considering worst case realization, worst case scenarios are considered in the constraints, possibly reducing the solution space and the optimality of the solution. However, notice that average performance can be

increased by using the previously mentioned scenario sampling methods with a larger probability of constraint violation [38], although we do not consider this possibility acceptable in the context of the current application.

V. Control Scheme Properties

In this section, it is proven that the design as shown in Figure 3 complies with the requirements of Section II-E. First, it is guaranteed that no crash occurs at all time, even when the platoon is subject to cyber-attack. Secondly, it is proven that, in healthy condition, there exists an input sequence such that the platoon conforms to the relaxed string stability as defined in Section II-A. Proofs of the presented Lemmas are presented in appendix A.

A. Safety Properties

To prove that crashes are avoided at all time, we will prove that safety Condition (17) always holds, even when the topology changes.

Lemma 3: Consider that at time instant k, a feasible solution of Problem (19) can be found by all vehicles $i \in \mathcal{N}$ satisfying Constraint (17) for $m = k+1$. Then, an input $u_i(k+1|k)$ exists such that Constraint (17) is also satisfied by all vehicles for $m = k+2$.

Lemma 4: Consider that Lemma 3 holds. Then, Constraint (17) with $m = k+2$ is also satisfied by all vehicles $i \in \mathcal{N}$ at time instant $k+1$, even when the communication topology changes.

Theorem 2: Using the MPC Controller (19) in each vehicle i with the topology switching rule as from Algorithm 1, it is guaranteed that $d_{i,s}(t) \geq 0 \ \forall i, t, s$.

Proof: In Lemmas 3 and 4 it has been shown that the safety constraints in the MPC Problem (19) are recursively feasible both for a constant topology and over topology switches. Furthermore, all safety constraints imply $d_{i,s}(t) \geq 0$, which proves the theorem statement. □

B. String Stability Properties

In this section, we will prove that in healthy conditions there always exists an input for which relaxed string stability (Definition 2) is achieved in the platoon. To this end, we will first prove that strict string stability can be achieved within each coalition. Then, it will be shown that the violation of the string stability between coalitions is bounded when using the proposed topology switching law, so that the whole platoon is relaxed string stable.

Lemma 5: If soft Constraint (18) holds for each vehicle $j \in \mathcal{C}$ with $\epsilon_s = 0$, there exist an input sequence $\mathbf{u}^i_{\mathcal{C}}$ for each vehicle $i \in \mathcal{C}$ such that the coalition is strictly string stable.

Theorem 3: Using the MPC Controller (19) and the topology switching law from Algorithm 1, there exist an input sequence $\mathbf{u}^i_{\mathcal{C}}$ for each vehicle $i \in \mathcal{C}$ such that the healthy CVP is relaxed string stability.

Proof: Lemma 5 proves strict string stability within a coalition. This only leaves us to prove that $v_j(k_2) - v_j(k_1) - (v_{j-1}(k_2) - v_{j-1}(k_1))$ is always upper-bounded between coalitions, i.e., for all $k_1 > 0, k_2 > k_1$, $j \in \mathcal{C}$, $j-1 \notin \mathcal{C}$. By the switching law from Algorithm 1

$$\begin{aligned} &v_j(k_2) - v_j(k_1) - (v_{j-1}(k_2) - v_{j-1}(k_1)) \\ &\quad = \Delta v_j(k_1) - \Delta v_j(k_2) \leq 2T_v \,, \end{aligned}$$

for all $k_1 > 0, k_2 > k_1$, $j \in \mathcal{N}$. □

Remark 11: The presented coalitional MPC method allows for safe control of CVPs subject to a large class of FDI as well as DoS cyber-attacks. Literature on cyber-attack tolerant MPC-based CVP control mostly addresses DoS attacks [28], [39], [40], [41], while work considering FDI attacks are limited [42], [43]. In [43], however, cyber-attacks are only mitigated when they cause violation of the safety constraints causing large tracking errors. Furthermore, [42] only allows for attacks of limited duration. ◁

TABLE II
Parameters Used in the Simulation Example

Parameter	Value	Parameter	Value
τ	$0.1\ [s^{-1}]$	Q_i	diag(100,0,0,0,10)
r	$10\ [m]$	R_i	50
h	$0.5\ [s]$	N_p	10
$u_{\max}$	$10\ [ms^{-2}]$	β_1	0.1
$u_{\min}$	$-10\ [ms^{-2}]$	β_2	1e5
T	$0.05\ [s]$	S_d	$\left\{ s \middle\vert \hat{u}_{p_{\mathcal{C}}} = 0, \begin{bmatrix} \hat{a}^i_{u_{\mathcal{C}}} \\ \hat{a}^{i,\text{ind}}_{y_{\mathcal{C}}} \end{bmatrix} = 0 \right\}$
T_v	$0.2\ [ms^{-1}]$	T_d	$0.2\ [m]$

VI. Simulation for Vehicle Platoon Control

In this section, the proposed control method is applied to a platoon of 4 vehicles following a leader vehicle. The input of the leader vehicle, which defines the platoon maneuvers is shown as the dashed line in Figure 4. The simulation parameters are given in Table II. The attacks injected in the communication are shown as dashed lines in Figure 6. As can be seen consecutively a step and a ramp attack are applied in the chosen scenario, which each target a different vehicle and measurement. Furthermore, the attacks are applied at times when the communication would nominally be active. This is done to show the versatility and effectiveness of the scheme.

Figure 4 shows the evolution of the states of all vehicles in this scenario. Figure 5 shows the evolution of the communication topology. Overall, the behaviour of the platoon is smooth and the tracking error over the whole scenario is at most 0.4m. Furthermore, note that when the tracking error is low, the platoon tends to operate in a decentralized manner as intended, thus saving coordination efforts. We would, however, like to shed some more light on a few noteworthy points.

First, one can see in Figure 4 that in the period between 1 and 2.2 seconds, the tracking error of vehicle 1 increases more than that of the other vehicles. This is because vehicle 1 cannot initiate communication with the preceding vehicle as this is the lead vehicle. This communication is beneficial especially while the platoon is decelerating as this causes the vehicles to have a negative relative velocity for reference tracking, which in turn causes the safety constraint (17) to become more restrictive. Vehicles 2, 3, and 4 limit the effect of the safety constraint by enabling communication with the

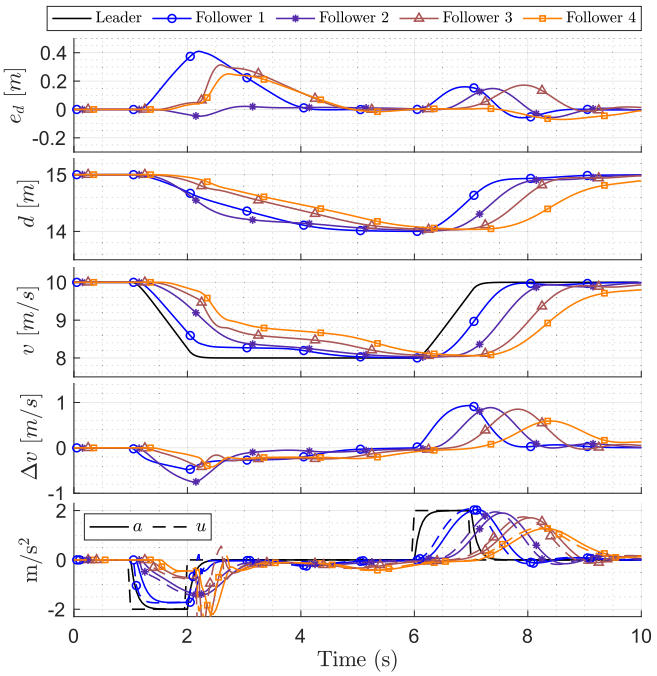

Fig. 4. Evolution of the states and input of all vehicles.

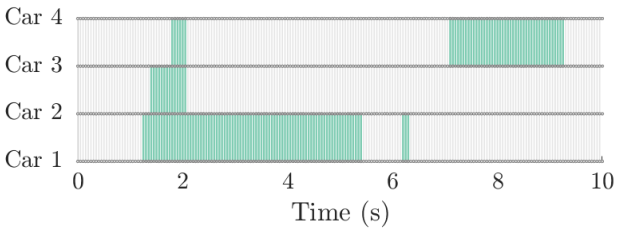

Fig. 5. Evolution of the communication topologies. Green lines indicate active communication links.

preceding vehicle (see Figure 5), which reduces uncertainty. However, vehicle 1 cannot initiate communication and therefore the safety constraint forces it to brake more than desired for reference tracking.

Notice also that vehicles 1 to 4 start operating with full communication from approximately 2s. Nevertheless, the detection of the attack on the signals that vehicle 3 transmits forces the isolation of vehicles 3 and 4 at around 2.2 $[s]$; hence the corresponding link is deactivated. The latter causes the spike in the acceleration of vehicles 3 and 4 which can be seen in Figure 4. At the time the communication with vehicle 4 is disabled, it is still decelerating and its relative velocity is negative. Therefore, to keep fulfilling the safety constraint after disabling communication, the relative velocity needs to be suddenly increased. Similarly, the communication with vehicle 2 is disabled when the attack on its communication is detected (see Figures 5 and 6).

Additionally, Table III compares the performance of the proposed coalitional control law with integrated cyber-attack detection with other possible communication topologies. Table III shows the cumulative costs obtained for different communication topologies and attacks scenarios. We use

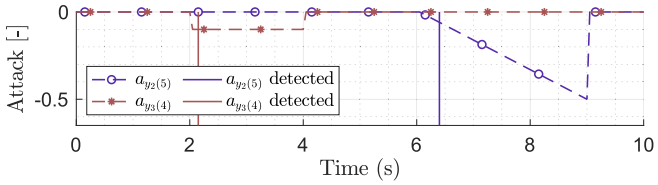

Fig. 6. Evolution of the attacks on the communication and the times of first detection.

TABLE III

VALUE OF THE CUMULATIVE COST FUNCTION FOR VARIOUS COMMUNICATION TOPOLOGIES AND ATTACK SCENARIOS

Attack scenario → Comm. topology ↓	None	Small (x2)	Medium (x3)	Large (x4)
Full comm.	9.60e3	9.40e3	1.18e4	Infeasible
No comm.	1.37e4	1.37e4	1.37e4	1.37e4
Switching	1.14e4	Infeasible	Infeasible	Infeasible
Switching with Attack Detection	1.14e4	1.27e4	1.27e4	1.27e4

"topology switching" to refer to that situation in which the vehicles change dynamically their communication topology only according to the tracking error and relative velocity, that is, the attacks are not considered and hence do not trigger the deactivation of any links. Also, "full communication" and "no communication" indicate, respectively, the situation in which all links are permanently enabled and disabled. The cumulative costs are provided for the case without attackers and for three other scenarios in which the severity of the attacks is progressively increased. In particular, we consider the output attacks shown in Figure 6 and scale them according to the factors indicated in Table III, while the lead vehicle follows the acceleration profile given in Figure 4. The attack magnitudes are chosen to highlight the differences between the communication topologies.

Firstly, one can see that the proposed *switching method with attack detection* incurs the same cost regardless of the attack magnitude. This is because the attacks are all quickly detected and the cost due to the attack is incurred during the mitigation which is equal for all attack magnitudes. This is different for the *switching method without attack detection*, which becomes infeasible, i.e. safety is no longer guaranteed, already for the smallest attack. This is because the attacked communication channels are not disabled. Likewise, the *full communication* method also becomes infeasible, but only for larger attacks. This is because with *full communication* there is more redundancy in measurements, which makes the method inherently more robust than the *switching* method.

Furthermore, it is interesting to note that the *full communication* method has lower costs than the proposed method for smaller attacks. This indicates that for small attacks the cost incurred by mitigation is larger than the cost incurred by the attack. Furthermore, one can see that the cost of the proposed method, for any attack magnitude, is smaller than the cost for the *no communication* method.

VII. CONCLUSION

A topology-switching coalitional MPC controller for collaborative vehicle platoons is introduced to guarantee safety

even when the communication between the vehicles is subject to cyber-attacks. To this end, the topology-switching coalitional MPC is integrated with a reduced order unknown input observer for cyber-attack detection.

The proposed MPC control law optimizes a cost function weighing performance and control effort to determine the control action subject to constraints that guarantee strict predecessor-follower string stability within each coalition. Likewise, a topology switching law enables/disables communication links when the tracking error or relative velocity between two vehicles exceeds/falls below a chosen threshold. This significantly reduces the cooperation costs with respect to a *full communication* approach. Furthermore, it allows the topology-switching coalitional MPC to obtain a relaxed string stability over the whole platoon.

In case of attack, the focus of the control law shifts towards guaranteeing safety. Firstly, when a cyber-attack is detected on a certain communication link, all communication with the transmitting vehicle is disabled. MPC constraints are in place to avoid crashes when such a forced topology switch occurs. Furthermore, even if the cyber-attacks are undetected safety is still guaranteed.

In summary, the designed control law provides a trade-off between performance and control effort while reducing the cooperation costs. Furthermore, relaxed string stability of the platoon can be obtained in nominal conditions, and safety is guaranteed even when the platoon is under attack. These properties have been shown theoretically and are illustrated using a platoon of 4 vehicles following a lead vehicle.

Future extensions will deal with more complex platoons, e.g., with vehicles leaving or entering the platoon, considering the lateral motion of the platoon or more realistic representations of the vehicle to vehicle communication.

APPENDIX

In what follows, we provide the proofs of Lemmas 3 to 5. For further use in these proofs, let us first introduce the following propositions:

Proposition 1: Variable $\delta(n|k)$ is positive if $u_i(\kappa) = u_{\min}$ $\forall k \leq \kappa < n$ and $\Delta v_{i,s}(n|k) < 0$.

Proof: By definition, $\gamma(k|k)$ represents an upper bound on the time to standstill when $u_i(\kappa) = u_{\min}$, $\forall \kappa \geq k$. Therefore, $\delta(n|k) = \gamma(k|k) - (n-k)T$ is an upper bound on the time left to standstill at instant $n > k$ if $u_i(\kappa) = u_{\min}$, $\forall k \leq \kappa < n$. Furthermore, $\Delta v_{i,s}(n|k) < 0$ implies that $v_{i,s}(n|k) > 0$, i.e. vehicle i is not at standstill at time n. This proves that, if $\Delta v_{i,s}(n|k) < 0$, then $\delta(n|k) > 0$. □

Proposition 2: The set $\{x_{i,s}(n|k)\}_{\forall s \in \mathcal{S}_e}$, which contains the predicted states of vehicle i at instant $(n|k)$ in the extreme scenarios, is bounded and form a convex set such that $x_{i,s}(n|k) \in \mathcal{X}_i^e(n|k) \triangleq \text{conv}(\{x_{i,s}(n|k)\}_{\forall s \in \mathcal{S}_e})$ $\forall s \in \mathcal{S}$.

Proof: We separately address the two sources of uncertainty. Firstly, the undetected cyber-attacks $\hat{a}^i_{y_{\mathcal{C}},s}$ are bounded and $\hat{a}^i_{y_{\mathcal{C}},s} \in \text{conv}(\{\hat{a}^i_{y_{\mathcal{C}},s}\}_{\forall s \in \mathcal{S}_e})$ $\forall s \in \mathcal{S}$. Furthermore, the realized state of each vehicle $i \in \mathcal{N}$ is always bounded. Therefore, $x_{\mathcal{C},s}(k|k)$ in Equation (14a) is also bounded and $x_{\mathcal{C},s}(k|k) \in \text{conv}(\{x_{\mathcal{C},s}(k|k)\}_{\forall s \in \mathcal{S}_e})$ $\forall s \in \mathcal{S}$.

Secondly, the unknown input of the vehicle $p_{\mathcal{C}}$, $\hat{u}_{p_{\mathcal{C}},s}$ is also bounded and $\hat{u}_{p_{\mathcal{C}},s} \in \text{conv}(\{\hat{u}_{p_{\mathcal{C}},s}\}_{\forall s \in \mathcal{S}_e})$, $\forall s \in \mathcal{S}$. As $w_{\mathcal{C},s} \sim a_{p_{\mathcal{C}},s}$ and $\text{conv}(\{a_{p_{\mathcal{C}},s}\}_{\forall s \in \mathcal{S}_e}) \subseteq \text{conv}(\{\hat{u}_{p_{\mathcal{C}},s}\}_{\forall s \in \mathcal{S}_e})$, we can also state that $w_{\mathcal{C},s}$ is bounded and $w_{\mathcal{C},s} \in \text{conv}(\{w_{\mathcal{C},s}\}_{\forall s \in \mathcal{S}_e})$, $\forall s \in \mathcal{S}$.

Therefore, Equation (14b) is a linear update equation for which $\mathcal{S}_e$ define convex bounds on both the input and the initial condition. As we consider $x_{i,s}(n|k)$ only for $n < k+N_p$, this is sufficient to prove $x_{i,s}(n|k)$ is bounded and $x_{i,s}(n|k) \in \text{conv}(\{x_{i,s}(n|k)\}_{\forall s \in \mathcal{S}_e})$ $\forall s \in \mathcal{S}$. □

Proof of Lemma 3: All Constraints (17) are lower bounds on $d_{i,s}(k+1|k)$ and each of the constraints is only active in a subset of the state-space, as discussed in Section IV-B. Therefore, we can consider each constraint separately. We will prove the lemma for each constraint by showing that if the constraint holds for $m = k+1$, using input $u_i(k+1|k) = u_{\min}$, it still holds for $m = k+2$. To this end, let us use the following relations:

$$d_{i,s}(k+2|k) = d_{i,s}(k+1|k) + T\Delta v_{i,s}(k+1|k), \tag{20a}$$

$$\Delta v_{i,s}(k+2|k) = \Delta v_{i,s}(k+1|k) + T\Delta a_{i,s}(k+1|k), \tag{20b}$$

$$\tau a_{i,s}(k+2|k) = (\tau - T)a_{i,s}(k+1|k) + Tu_i(k+1|k). \tag{20c}$$

Now, firstly, consider Constraint (17a) for $d_{i,s}(k+1|k)$, which is active only if $\Delta v_i(k+1|k) \geq 0$.[3] Then, if Constraint (17a) is satisfied, i.e., $d_{i,s}(k+1|k) \geq 0$, the following holds:

$$d_{i,s}(k+2|k) = d_{i,s}(k+1|k) + T\Delta v_{i,s}(k+1|k) \geq 0,$$

which proves the lemma for Constraint (17a).

Secondly, consider Constraint (17b), which is active only if $\Delta v_{i,s}(k+1|k) \leq 0$ and $\Delta a_{i,s}(k+1|k) \geq 0$, and recall Proposition 1. Then, if Constraint (17b) is satisfied, i.e. $d_{i,s}(k+1|k) \geq -\Delta v_{i,s}(k+1|k)\delta(k+1|k)$, the following holds:

$$\begin{aligned} d_{i,s}(k+2|k) &= d_{i,s}(k+1|k) + T\Delta v_{i,s}(k+1|k) \\ &\geq -\Delta v_{i,s}(k+1|k)\delta(k+1|k) + T\Delta v_{i,s}(k+1|k) \\ &= -\Delta v_{i,s}(k+1|k)\delta(k+2|k) \\ &= -(\Delta v_{i,s}(k+2|k) - \Delta a_{i,s}(k+1|k)T)\delta(k+2|k) \\ &\geq -\Delta v_{i,s}(k+2|k)\delta(k+2|k), \end{aligned}$$

where we have used (20a) and (20b), and the fact that $\Delta a_{i,s}(k+1|k) \geq 0$ in the last inequality. This proves the lemma for Constraint (17b).

Lastly, consider Constraint (17c), which is active only if $\Delta a_{i,s}(k+1|k) \leq 0$. Then, if Constraint (17c) is satisfied,

[3]Note that this is a necessary condition for Constraint (17a) to be active. If $\Delta v_{i,s}(k+1|k) < 0$, Constraint (17b) is more restrictive. However, if $\tau\Delta a_i(k+1|k) < -\Delta v_i(k+1|k)$ Constraint (17c) can be more restrictive even if $\Delta v_i(k+1|k) \geq 0$.

we can use (20) to derive the following:

$$\begin{aligned}
&d_{i,s}(k+2|k)\\
&\quad= d_{i,s}(k+1|k) + T\Delta v_{i,s}(k+1|k)\\
&\quad\geq -(\Delta v_{i,s}(k+1|k) + \tau\Delta a_{i,s}(k+1|k))\delta(k+1|k)\\
&\qquad + T\Delta v_{i,s}(k+1|k)\\
&\quad\geq -(\Delta v_{i,s}(k+1|k) + \tau\Delta a_{i,s}(k+1|k))\delta(k+1|k)\\
&\qquad + T(\Delta v_{i,s}(k+1|k) + \tau\Delta a_{i,s}(k+1|k))\\
&\quad= -(\Delta v_{i,s}(k+1|k) + \tau\Delta a_{i,s}(k+1|k))\delta(k+2|k)\\
&\quad\geq -(\Delta v_{i,s}(k+2|k) - T\Delta a_{i,s}(k+1|k))\delta(k+2|k)\\
&\qquad - (\tau\Delta a_{i,s}(k+2|k) + T\Delta a_{i,s}(k+1|k))\delta(k+2|k)\\
&\qquad + T\Delta u_{i,s}(k+1|k)\delta(k+2|k),
\end{aligned}$$

where $\Delta u_{i,s}(k+1|k) \triangleq u_{i-1,s}(k+1|k) - u_{i,s}(k+1|k)$, such that with the chosen input, $\Delta u_i(k+1|k) = u_{i-1}(k+1|k) - u_{\min} \geq 0$. Then,

$$\begin{aligned}
&d_{i,s}(k+2|k) \geq\\
&\quad -(\Delta v_{i,s}(k+2|k) + \tau\Delta a_{i,s}(k+2|k))\delta(k+2|k),
\end{aligned}$$

which proves the lemma for Constraint (17c).[4]

Proof of Lemma 4: Without loss of generality, consider a only the relation between vehicles $i, i-1 \in \mathcal{N}$ and the following changes to the coalitions:

(a) At instant k, the vehicles form a coalition $\mathcal{C} = \{i-1, i\}$, and it *breaks up* into $\mathcal{C}_1 = \{i-1\}$ and $\mathcal{C}_2 = \{i\}$ at $k+1$.
(b) At instant k, the vehicles are in different coalitions, say $\mathcal{C}_1 = \{i-1\}$ and $\mathcal{C}_2 = \{i\}$, and they *join* into a single $\mathcal{C} = \{i-1, i\}$ at $k+1$.
(c) Coalition $\mathcal{C} = \{i-1, i\}$ remains constant.

In Lemma 3, it is proven that if (17) holds for $x_{i,s}(k+1|k)$, there exists an input sequence such that it also holds for $x_{i,s}(k+2|k),\ \forall s \in \mathcal{S}$. Therefore, a sufficient condition for Constraint (17) to hold also for $x_{i,s}(k+2|k+1)$ is

$$x_{i,s}(k+2|k+1) \in \mathcal{X}_i^{\mathrm{e}}(k+2|k),\ \forall k \geq 0,\ \forall s \in \mathcal{S}. \tag{21}$$

The existence of set $\mathcal{X}_i^{\mathrm{e}}(k+2|k)$ is proven by Proposition 2. Using the prediction model in (14), we have

$$\begin{aligned}
x_{i,s}(k+2|k+1) &= A_i x_{i,s}(k+1|k+1) + B_i u_i(k+1|k+1)\\
&\quad + A_i^w x_{i-1,s}(k+1|k+1),\\
x_{i,s}(k+2|k) &= A_i x_{i,s}(k+1|k) + B_i u_i(k+1|k)\\
&\quad + A_i^w x_{i-1,s}(k+1|k),
\end{aligned} \tag{22}$$

where, without loss of generality, we choose $u_i(k+1|k+1) = u_i(k+1|k)$. Furthermore, by Proposition 2,

$$x_{i,s}(k+1|k) \in \mathcal{X}_i^{\mathrm{e}}(k+1|k),\ \forall i \in \mathcal{C},\ \forall s \in \mathcal{S},\ \forall k \geq 0,$$

and as the realised state $x_{i,s}(k+1|k+1)$ is the outcome of one of the possible scenarios in $\mathcal{S}$, we also have

$$x_{i,s}(k+1|k+1) \in \mathcal{X}_i^{\mathrm{e}}(k+1|k),\ \forall i \in \mathcal{C},\ \forall s \in \mathcal{S},\ \forall k \geq 0. \tag{23}$$

[4]Above the lemma is proved if the active constraint is fixed. It is, however, possible that the active constraint changes between time instants $k+1$ and $k+2$. Similar approaches can be used to prove the lemma for each of these cases. For brevity, however, these full proofs are omitted.

That is, the new initial condition $x_{i,s}(k+1|k+1)$ is bounded by the prediction based on the extreme scenario at time k.

Consider case (a), where using Equation (14d), $x_{i-1,s}(k+1|k+1) = x_{i-1,s}(k+1|k) \in \mathcal{X}_{i-1}^{\mathrm{e}}(k+1|k)$ for all $s \in \mathcal{S}$. In cases (b) and (c), as vehicles i and $i-1$ are in the same coalition at time $k+1$, we can directly apply Equation (23) for vehicle $i-1$ too, such that $x_{i-1,s}(k+1|k+1) \in \mathcal{X}_{i-1}^{\mathrm{e}}(k+1|k)$ for all $s \in \mathcal{S}$.

Substituting the results above into Equation (22) implies Equation (21) holds, proving the lemma.

Proof of Lemma 5: Define $dv_i^{n_2}(n_1) \triangleq v_i(n_1+n_2) - v_i(n_1)$ as the change in velocity of vehicle i over a period of n_2 time-steps, and recall $dv_{i,s}(k) = dv_{i,s}^{\mathrm{pos}}(k) + dv_{i,s}^{\mathrm{neg}}(k) = v_{i,s}(k+N_{\mathrm{p}}|k) - v_{i,s}(k|k)$ is the predicted change in velocity over the length of the prediction horizon.

Now, without loss of generality, consider that, dictated by the considered maneuver of the lead vehicle, $dv_{i,s}(k) \geq 0,\ \forall k \in [0, N)$ and $dv_{i,s}(k) \leq 0,\ \forall k \in [N, N_2)$. Furthermore, consider initially $\Delta v_i(0) \leq 0$. Starting from this initial condition, by constraint (18c)

$$\exists\ n_s < N_{\mathrm{p}} \text{ s.t. } \Delta v_{i,s}(n_s) \geq 0 \quad \Delta v_{i,s}(k) \leq 0,\ \forall k \in [0, n_s), \tag{24}$$

for all $s \in \mathcal{S}_0$. Then, by the definition of the scenarios $\mathcal{S}_0$, also

$$\exists\ n < N_{\mathrm{p}} \text{ s.t. } \Delta v_i(n) \geq 0\ \&\ \Delta v_i(k) \leq 0,\ \forall k \in [0, n), \tag{25}$$

Furthermore,

$$\exists\ \mathbf{u}_i \text{ s.t. } \Delta v_i(k) \leq 0,\ \forall k \in (n, N). \tag{26}$$

Note that $u_i(k) = u_{i-1}(k)\ \forall k \in (n, N)$ is one of the input sequences that guarantees this. With this we can derive

$$dv_i^{\ell-k}(k) - dv_{i-1}^{\ell-k}(k) = \Delta v_i(k) - \Delta v_i(\ell) \leq 0,\ \forall k \in [0, \ell), \tag{27}$$

where $\ell < N$. This implies strict string stability according to Definition 1 for time steps 0 to N.

At time N we have $\Delta v_i(N) \geq 0$ and $dv_{i,s}(N) \leq 0$, which is a similar situation to the the initial one. Therefore, following the same line of reasoning, $\exists\ n_2 < N + N_{\mathrm{p}}$ such that

$$\Delta v_i(n_2) \leq 0\ \&\ \Delta v_i(k) \geq 0,\ \forall k \in [N, N+n_2), \tag{28}$$

and

$$\exists\ \mathbf{u}_i \text{ s.t. } \Delta v_i(k) \geq 0,\ \forall k \in (N+n_2, N_2), \tag{29}$$

such that

$$dv_i^{\ell-k}(k) - dv_{i-1}^{\ell-k}(k) = \Delta v_i(k) - \Delta v_i(\ell) \geq 0, \tag{30}$$

for all $k \geq N$ and $k < N + n_2 \leq \ell < N_2$, which implies strict string stability for time steps N to N_2. At time step N_2 the situation is then as it was initially, such that the proof can be repeated for all time steps.

REFERENCES

[1] J. Ploeg, B. T. M. Scheepers, E. van Nunen, N. van de Wouw, and H. Nijmeijer, "Design and experimental evaluation of cooperative adaptive cruise control," in *Proc. 14th Int. IEEE Conf. Intell. Transp. Syst. (ITSC)*, Oct. 2011, pp. 260–265.

[2] T. van der Sande and H. Nijmeijer, "From cooperative to autonomous vehicles," in *Sensing and Control for Autonomous Vehicles: Applications to Land, Water and Air Vehicles*. Cham, Switzerland: Springer, 2017, pp. 435–452.

[3] S. E. Li, Y. Zheng, K. Li, L.-Y. Wang, and H. Zhang, "Platoon control of connected vehicles from a networked control perspective: Literature review, component modeling, and controller synthesis," *IEEE Trans. Veh. Technol.*, early access, Jul. 6, 2017, doi: 10.1109/TVT.2017.2723881.

[4] L. D. Baskar, B. De Schutter, J. Hellendoorn, and Z. Papp, "Traffic control and intelligent vehicle highway systems: A survey," *IET Intell. Transp. Syst.*, vol. 5, no. 1, pp. 38–52, Mar. 2011.

[5] L. Y. Wang, A. Syed, G. G. Yin, A. Pandya, and H. Zhang, "Control of vehicle platoons for highway safety and efficient utility: Consensus with communications and vehicle dynamics," *J. Syst. Sci. Complex.*, vol. 27, no. 4, pp. 605–631, Aug. 2014.

[6] S. Feng, H. Sun, Y. Zhang, J. Zheng, H. X. Liu, and L. Li, "Tube-based discrete controller design for vehicle platoons subject to disturbances and saturation constraints," *IEEE Trans. Control Syst. Technol.*, vol. 28, no. 3, pp. 1066–1073, May 2020.

[7] S. Feng, Z. Song, Z. Li, Y. Zhang, and L. Li, "Robust platoon control in mixed traffic flow based on tube model predictive control," *IEEE Trans. Intell. Vehicles*, vol. 6, no. 4, pp. 711–722, Dec. 2021.

[8] Y. Zheng, S. E. Li, K. Li, F. Borrelli, and J. K. Hedrick, "Distributed model predictive control for heterogeneous vehicle platoons under unidirectional topologies," *IEEE Trans. Control Syst. Technol.*, vol. 25, no. 3, pp. 899–910, May 2017.

[9] W. B. Dunbar and D. S. Caveney, "Distributed receding horizon control of vehicle platoons: Stability and string stability," *IEEE Trans. Autom. Control*, vol. 57, no. 3, pp. 620–633, Mar. 2012.

[10] E. Camponogara, D. Jia, B. Krogh, and S. Talukdar, "Distributed model predictive control," *IEEE Control Syst. Mag.*, vol. 22, no. 1, pp. 44–52, Feb. 2002.

[11] F. Fele, J. M. Maestre, and E. F. Camacho, "Coalitional control: Cooperative game theory and control," *IEEE Control Syst. Mag.*, vol. 37, no. 1, pp. 53–69, Feb. 2017.

[12] P. R. Baldivieso-Monasterios and P. A. Trodden, "Coalitional predictive control: Consensus-based coalition forming with robust regulation," *Automatica*, vol. 125, Mar. 2021, Art. no. 109380.

[13] A. Maxim and C.-F. Caruntu, "Coalitional distributed model predictive control strategy for vehicle platooning applications," *Sensors*, vol. 22, no. 3, p. 997, Jan. 2022. [Online]. Available: https://www.mdpi.com/1424-8220/22/3/997

[14] P. Chanfreut, J. M. Maestre, and E. F. Camacho, "A survey on clustering methods for distributed and networked control systems," *Annu. Rev. Control*, vol. 52, pp. 75–90, Jan. 2021.

[15] F. Fele, J. M. Maestre, S. M. Hashemy, D. M. de la Peña, and E. F. Camacho, "Coalitional model predictive control of an irrigation canal," *J. Process Control*, vol. 24, no. 4, pp. 314–325, Apr. 2014.

[16] P. Chanfreut, J. M. Maestre, and E. F. Camacho, "Coalitional model predictive control on freeways traffic networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 11, pp. 6772–6783, Nov. 2021.

[17] E. Masero, J. R. D. Frejo, J. M. Maestre, and E. F. Camacho, "A light clustering model predictive control approach to maximize thermal power in solar parabolic-trough plants," *Sol. Energy*, vol. 214, pp. 531–541, Jan. 2021.

[18] K. Li, Y. Bian, S. E. Li, B. Xu, and J. Wang, "Distributed model predictive control of multi-vehicle systems with switching communication topologies," *Transp. Res. C, Emerg. Technol.*, vol. 118, Sep. 2020, Art. no. 102717.

[19] B. Ding, L. Ge, H. Pan, and P. Wang, "Distributed MPC for tracking and formation of homogeneous multi-agent system with time-varying communication topology," *Asian J. Control*, vol. 18, no. 3, pp. 1030–1041, May 2016.

[20] P. Wang, H. Deng, J. Zhang, L. Wang, M. Zhang, and Y. Li, "Model predictive control for connected vehicle platoon under switching communication topology," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 7, pp. 7817–7830, Jul. 2022.

[21] H. Zhao, X. Dai, Q. Zhang, and J. Ding, "Robust event-triggered model predictive control for multiple high-speed trains with switching topologies," *IEEE Trans. Veh. Technol.*, vol. 69, no. 5, pp. 4700–4710, May 2020.

[22] Y. A. Harfouch, S. Yuan, and S. Baldi, "An adaptive switched control approach to heterogeneous platooning with intervehicle communication losses," *IEEE Trans. Control Netw. Syst.*, vol. 5, no. 3, pp. 1434–1444, Sep. 2018.

[23] Y. Zheng, S. E. Li, J. Wang, D. Cao, and K. Li, "Stability and scalability of homogeneous vehicular platoon: Study on the influence of information flow topologies," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 1, pp. 14–26, Jan. 2015.

[24] J. J. Haas, "The effects of wireless jamming on vehicle platooning," Univ. Illinois, Champaign, IL, USA, Tech. Rep., 2009. Accessed: Dec. 8, 2021. [Online]. Available: https://www.wiki.illinois.edu/wiki/download/attachments/360546327/jjhaas2_platoon_jamming.pdf

[25] M. Amoozadeh et al., "Security vulnerabilities of connected vehicle streams and their impact on cooperative driving," *IEEE Commun. Mag.*, vol. 53, no. 6, pp. 126–132, Jun. 2015.

[26] R. van der Heijden, T. Lukaseder, and F. Kargl, "Analyzing attacks on cooperative adaptive cruise control (CACC)," in *Proc. IEEE Veh. Netw. Conf. (VNC)*, Nov. 2017, pp. 45–52.

[27] Z. A. Biron, S. Dey, and P. Pisu, "Resilient control strategy under denial of service in connected vehicles," in *Proc. Amer. Control Conf. (ACC)*, May 2017, pp. 4971–4976.

[28] M. H. Basiri, N. L. Azad, and S. Fischmeister, "Attack resilient heterogeneous vehicle platooning using secure distributed nonlinear model predictive control," in *Proc. 28th Medit. Conf. Control Autom. (MED)*, Sep. 2020, pp. 307–312.

[29] E. van Nunen, J. Ploeg, A. M. Medina, and H. Nijmeijer, "Fault tolerancy in cooperative adaptive cruise control," in *Proc. 16th Int. IEEE Conf. Intell. Transp. Syst. (ITSC)*, Oct. 2013, pp. 1184–1189.

[30] A. Petrillo, A. Pescapé, and S. Santini, "A collaborative approach for improving the security of vehicular scenarios: The case of platooning," *Comput. Commun.*, vol. 122, pp. 59–75, Jun. 2018.

[31] J. Lan and R. J. Patton, "A new strategy for integration of fault estimation within fault-tolerant control," *Automatica*, vol. 69, pp. 48–59, Jul. 2016.

[32] X. Feng and R. Patton, "Active fault tolerant control of a wind turbine via fuzzy MPC and moving horizon estimation," *IFAC Proc. Volumes*, vol. 47, no. 3, pp. 3633–3638, 2014, doi: 10.3182/20140824-6-ZA-1003.00965.

[33] S. Tong, B. Huo, and Y. Li, "Observer-based adaptive decentralized fuzzy fault-tolerant control of nonlinear large-scale systems with actuator failures," *IEEE Trans. Fuzzy Syst.*, vol. 22, no. 1, pp. 1–15, Feb. 2014.

[34] R. Hmidi, A. B. Brahim, F. B. Hmida, and A. Sellami, "Robust fault tolerant control design for nonlinear systems not satisfying matching and minimum phase conditions," *Int. J. Control, Autom. Syst.*, vol. 18, no. 9, pp. 2206–2219, Sep. 2020.

[35] J. Ploeg, D. P. Shukla, N. van de Wouw, and H. Nijmeijer, "Controller synthesis for string stability of vehicle platoons," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 2, pp. 854–865, Apr. 2014.

[36] J. M. Maestre, D. M. de la Peña, A. J. Losada, E. Algaba, and E. F. Camacho, "A coalitional control scheme with applications to cooperative game theory," *Optim. Control Appl. Methods*, vol. 35, no. 5, pp. 592–608, Sep. 2014.

[37] R. M. G. Ferrari and A. M. H. Teixeira, "A switching multiplicative watermarking scheme for detection of stealthy cyber-attacks," *IEEE Trans. Autom. Control*, vol. 66, no. 6, pp. 2558–2573, Jun. 2021.

[38] M. C. Campi, S. Garatti, and M. Prandini, "The scenario approach for systems and control design," *Annu. Rev. Control*, vol. 33, no. 2, pp. 149–157, 2009.

[39] M. Basiri, N. Azad, and S. Fischmeister, "Secure dynamic nonlinear heterogeneous vehicle platooning: Denial-of-service cyber-attack case," in *Security in Cyber-Physical Systems*. Cham, Switzerland: Springer, 2021, pp. 287–315.

[40] H. Sun, C. Peng, and F. Ding, "Self-discipline predictive control of autonomous vehicles against denial of service attacks," *Asian J. Control*, vol. 24, no. 6, pp. 3538–3551, Nov. 2022.

[41] J. Chen, H. Zhang, and G. Yin, "Distributed dynamic event-triggered secure model predictive control of vehicle platoon against DoS attacks," *IEEE Trans. Veh. Technol.*, vol. 72, no. 3, pp. 2863–2877, Mar. 2023.

[42] G. Franzè, F. Tedesco, and D. Famularo, "A distributed resilient control strategy for leader-follower systems under replay attacks," in *Proc. 7th Int. Conf. Control, Decis. Inf. Technol. (CoDIT)*, vol. 1, Jun. 2020, pp. 469–474.

[43] H. Lyu, T. Wang, R. Cheng, and H. Ge, "Improved longitudinal control strategy for connected and automated truck platoon against cyberattacks," *IET Intell. Transp. Syst.*, vol. 16, no. 12, pp. 1710–1725, Dec. 2022.

Twan Keijzer received the M.Sc. degree in aerospace engineering from Delft University of Technology, The Netherlands, in 2018, and the Ph.D. degree from Delft Center for Systems and Control, in 2023, under the supervision of Riccardo Ferrari. He is currently an Research and Development Engineer with the Royal Netherlands Aerospace Centre. His research is mainly applied to the aerospace and automotive sector and focuses on fault and cyber attack detection methods, distributed fault tolerant control approaches, and experimental validation.

Paula Chanfreut received the Ph.D. degree in automation engineering from the University of Seville, Spain, in 2022. She was a Predoctoral Fellow with the University of Seville, under the Spanish University Professor Training Program (FPU). From 2022 to 2023, she was with ERC Advanced Grant OCONTSOLAR. She is currently an Assistant Professor with the Department of Mechanical Engineering, Eindhoven University of Technology, The Netherlands. Her research is framed within the field of MPC, with emphasis on its non-centralized implementations.

José María Maestre (Senior Member, IEEE) received the Ph.D. degree from the University of Seville. He has held various positions at universities, such as TU Delft, University of Pavia, University of Kyoto, and Tokyo Institute of Technology. He is currently a Full Professor with the University of Seville. He has published over 200 journal and conference papers, co-edited several books, and led several research projects. His research interests include the control of distributed cyber-physical systems, with a special emphasis on the integration of heterogeneous agents in the control loop. Finally, his achievements have been recognized through several awards and honors, including Spanish Royal Academy of Engineering's Medal for his contributions to the predictive control of large-scale systems.

Riccardo Maria Giorgio Ferrari (Member, IEEE) received the Laurea degree (cum laude and printing honours) in electronic engineering and the Ph.D. degree in information engineering from the University of Trieste, Italy, in 2004 and 2009, respectively. He hold both academic and industrial research and development positions, in particular as a Researcher in the field of process instrumentation and control for the steel-making sector. He is a Marie Curie alumnus and currently an Associate Professor with Delft Center for Systems and Control, Delft University of Technology, The Netherlands. His research interests include wind power fault tolerant control and fault diagnosis and attack detection in large-scale cyber-physical systems, with applications to electric vehicles, cooperative autonomous vehicles, and industrial control systems. He was a recipient of the 2005 Giacomini Award of Italian Acoustic Society and he obtained the 2nd place in the Competition on Fault Detection and Fault Tolerant Control for Wind Turbines during IFAC 2011. Furthermore, he was awarded an Honorable Mention for the Pauk M. Frank Award at the IFAC SAFEPROCESS in 2018 and won an Airbus Award at IFAC 2020 for the best contribution to the competition on Aerospace Industrial Fault Detection.