Detection of Cyber-Attacks in Collaborative Intersection Control

Keijzer, Twan; Jarmolowitz, Fabian; Ferrari, Riccardo M.G.

DOI
10.23919/ECC54610.2021.9655088

Publication date
2021

Document Version
Final published version

Published in
Proceedings of the European Control Conference (ECC 2021)

2021 European Control Conference (ECC)
June 29 - July 2, 2021. Rotterdam, Netherlands

Detection of Cyber-Attacks in Collaborative Intersection Control

Twan Keijzer[1], Fabian Jarmolowitz[2], and Riccardo M.G. Ferrari[1]

Abstract— **Road intersections are widely recognized as a lead cause for accidents and traffic delays. In a future scenario with a significant adoption of Cooperative Autonomous Vehicles, solutions based on fully automatic, signage-less Intersection Control would become viable. Such a solution, however, requires communication between vehicles and, possibly, the infrastructure over wireless networks. This increases the attack surface available to a malicious actor, which could lead to dangerous situations. In this paper, we address the safety of Intersection Control algorithms, and design a Sliding-Mode-Observer based solution capable of detecting and estimating false data injection attacks affecting vehicles' communication. With respect to previous literature, a novel detection logic with improved detection performances is presented. Simulation results are provided to show the effectiveness of the proposed approach.**

I. Introduction

Classical road intersections for human-driven vehicles are managed via fixed signage and traffic lights, which allows for a sub-optimal vehicle throughput while guaranteeing an adequate safety level. Still, it is well known that classical intersections are a major cause of accidents, due to human error, and traffic delays [1]. In the early 2000s, based on the projected introduction of *Cooperative Autonomous Vehicles* (CAV), works such as [2] proposed replacing classical intersections with safer, automated solutions. *Intersection Control* (IC) would thus automatize the tasks of negotiating, planning and executing the trajectories of CAVs in order to increase the safety and vehicle throughput of the junction.

A key enabling technology of an IC solution are wireless networks, allowing real-time communication of measurements and control signals between CAVs, and to the road-side infrastructure. The security and robustness of such Vehicle-to-Everything (V2X) networks, is thus of paramount importance for the safety of the IC itself. As a first line of defence, current V2X protocols include encryption and authentication mechanisms to prevent intrusion. For instance, the Autosar standard with End-to-End Protection (E2E), following the ISO 26262 standard, is such a preventive protection measure and is analyzed in [3]. The case of communication disruption or false data injection by a malicious attacker [4], [5] that can circumvent these protections would instead require a different approach. Indeed, a major difference between an ordinary fault and a smart attacker is that the latter actively seeks to cause great harm to the IC, while minimizing the possibility of being detected [6], [7]. This lead to development of increasingly sophisticated anomaly detection approaches for V2X communication traffic, as a second line of defence that can provide a guarantee against either inside attackers, or sophisticated attackers that successfully infiltrated the system. These techniques include plausibility checks based on elementary models of the CAVs and IC [8], [9], [10], as well as more advanced, model-based approaches that have been proposed in the literature for detecting attacks in general Cyber-Physical Systems (CPS), such as [11], [12], [13], [14], [15]. Nevertheless, these techniques were not yet applied to CAVs on an automated intersection.

In this paper, we will address the problem of designing a second line of defence for IC subjected to false-data injection attacks. To this end, a novel cyber-attack detection method is presented based on a *Sliding Mode Observer* (SMO). In previous work by the authors [14] a detection logic based on the so-called *Equivalent Output Injection* (EOI) term of the SMO was presented. The novel approach proposed in the present paper no longer requires this EOI, leading to better detection performance.The proposed technique is verified in simulation of an IC scenario. Here a control approach called *Virtual Platooning* (VP, [16]) is used. However, also potentially better-performing optimization-based approaches could be applied [17], [18], [19], [20] without affecting the proposed detection method.

The structure of the paper is as follows: Section II introduces the problem addressed in this paper, while Sections III and IV, respectively, introduce the SMO design and the detection thresholds on which the proposed attack detection strategy is built. Simulation results are shown in Section V, while concluding remarks are finally drawn in Section VI.

II. Problem Formulation

In this paper, detection of cyber-attacks on the inter-vehicle communication is considered in a collaborative intersection control scenario. Each car is modeled as

$$\left\{\begin{aligned} \begin{bmatrix}\dot{p}_i\\ \dot{v}_i\\ \dot{a}_i\end{bmatrix} &= \underbrace{\begin{bmatrix}0 & 1 & 0\\ 0 & 0 & 1\\ 0 & 0 & -\frac{1}{\tau_i}\end{bmatrix}}_{A_i}\underbrace{\begin{bmatrix}p_i\\ v_i\\ a_i\end{bmatrix}}_{x_i} + \underbrace{\begin{bmatrix}0\\ 0\\ \frac{1}{\tau_i}\end{bmatrix}}_{B_i} u_i\,,\\ y_i &= \begin{bmatrix}p_i - p_{i-1} - L_i\\ v_i - v_{i-1}\\ v_i\\ a_i\end{bmatrix} + \zeta_i\,,\end{aligned}\right. \tag{1}$$

where the subscripts i and $i-1$ denote the variables are related to cars i and $i-1$ respectively. p, v, a, u, y, ζ, τ, and L are, respectively, the distance from the rear of the vehicle to the intersection (negative when approaching the intersection), velocity, acceleration, input, measurements, sensor noise, engine time constant, and length of the cars.

The considered IC scenario is depicted in Figure 1. In this scenario, as can be seen from Equation (1), each car measures the relative distance to the intersection, and relative velocity from the preceding car, as well as its own velocity and acceleration. Mandatory for having these

[1] Twan Keijzer and Riccardo M.G. Ferrari are with Delft Centre for Systems and Control, Delft University of Technology, 2628 CD Delft, The Netherlands {t.keijzer,r.ferrari}@tudelft.nl

[2] Fabian Jarmolowitz is with Corporate Sector Research and Advanced Engineering, Robert-Bosch GmbH, 71272 Renningen, Germany fabian.jarmolowitz@bosch.com

978-94-6384-236-5 ©2021 EUCA

measurements for all cars entering the intersection are either a central infrastructure with appropriate sensors and V2X communication and/or cars equipped with lateral sensors, e.g. [21]. Furthermore, each car receives, via wireless V2V communication, the input of the preceding car.

The interaction between two cars in an IC scenario can be modeled, from the perspective of car i, as

$$\begin{cases} \begin{bmatrix} \dot{x}_{i-1} \\ \dot{x}_i \end{bmatrix} = \underbrace{\begin{bmatrix} A_{i-1} & 0 \\ 0 & A_i \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_{i-1} \\ x_i \end{bmatrix}}_{x} + \underbrace{\begin{bmatrix} B_{i-1} & 0 \\ 0 & B_i \end{bmatrix}}_{B} \underbrace{\begin{bmatrix} u_{i-1} \\ u_i \end{bmatrix}}_{u}, \\ y_i = C \underbrace{\begin{bmatrix} x_{i-1} \\ x_i \end{bmatrix}}_{x} + \underbrace{\begin{bmatrix} -L_i \\ 0_{3\times 1} \end{bmatrix}}_{c} + \zeta_i \,. \end{cases} \tag{2}$$

Here, C can be derived from equation (1). Furthermore, one can see that the coupling between the vehicles appears in the measurement equation only. These measurements, as well as the communicated input u_{i-1}, can be used by car i to calculate a control input u_i such that the IC objective is achieved. In this work, which primarily deals with the cyber-attack detection, any control law for IC can be chosen without affecting the detection method.

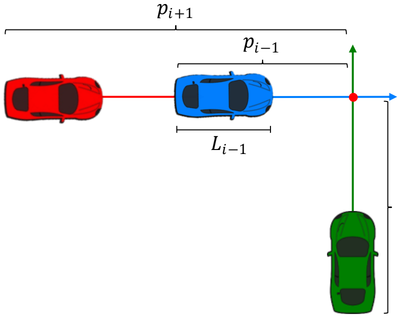

Fig. 1. Intersection Control Scenario [22]

A. Model Uncertainty & Cyber-attack

In this section, System (2) will be rewritten to make the model uncertainty and cyber-attack explicit. To quantify the model uncertainty, the following assumption is made.

Assumption 1: Each car i is assumed know its own dynamics, represented by τ_i and L_i. It will, however, only have access to the nominal dynamics of the preceding car.

The nominal value of τ_{i-1}, available to car i, is defined as

$$\hat{\tau}_{i-1} = r_\tau \tau_{i-1} \,.$$

Furthermore, the cyber-attack is defined as a "man in the middle" attack, such that each car i will receive

$$u_{i-1,r} = u_{i-1} + \Delta u_{i-1} \,,$$

where Δu_{i-1} is the cyber-attack. To make the model uncertainty and cyber-attack explicit a model of the form

$$\begin{cases} \dot{x} = Ax + Bu + E\eta + F\Delta u_{i-1} \,, \\ y_i = Cx + c + \zeta_i \,, \end{cases} \tag{3}$$

is proposed. Here A, B, and u are redefined using

$$\tau_{i-1} \leftarrow \hat{\tau}_{i-1} \text{ and } u_{i-1} \leftarrow u_{i-1,r} \,,$$

such that they are known to car i, i.e. they can be used in its detection logic. Furthermore, E, F, and η are defined as

$$E = \begin{bmatrix} 0_{2\times 1} \\ \frac{1}{\hat{\tau}_{i-1}} \\ 0_{3\times 1} \end{bmatrix} ; F = \begin{bmatrix} 0_{2\times 1} \\ -\frac{1}{\hat{\tau}_{i-1}} \\ 0_{3\times 1} \end{bmatrix} ; \eta = (r_\tau - 1)(u_{i-1} - a_{i-1}) \,,$$

where E and F are known to car i and η is unknown uncertainty. Note that Systems (3) and (2) are only reformulated.

The following assumptions are made on system (3).

Assumption 2: The sensor noise ζ_i is zero-mean and bounded by a known value $\bar{\zeta}_i \geq |\zeta_i|$.

Assumption 3: The uncertainty η, and cyber-attack Δu_{i-1} are bounded by known values $\bar{\eta} \geq |\eta|$ and $\bar{\Delta} \geq |\Delta u_{i-1}|$. These bounds are defined for the IC scenario in section V.

B. Model Transformation

In System (3) the cyber-attack appears as an unknown input. This allows for the use of an SMO for estimation and detection of the cyber-attack [14]. In order to implement this SMO based approach, the system is transformed to

$$\begin{cases} \dot{x}_1 = A_{11}x_1 + A_{12}x_2 + B_1 u + E_1\eta + F_1\Delta u_{i-1} \\ \dot{x}_2 = A_{21}x_1 + A_{22}x_2 + B_2 u + E_2\eta + F_2\Delta u_{i-1} \\ y_i = x_2 + c + \zeta_i \end{cases} \tag{4}$$

using a transformation introduced by [23]. Here it is required that $A_{11} \prec 0$ to ensure that the observer error dynamics are stable. The following assumption ensures $A_{11} \prec 0$ [23].

Assumption 4: The invariant zeros of (A,F,C) lie in $\mathbb{C}_{-}$.

III. Sliding Mode Observer Design

Based on [14], the following SMO is introduced

$$\begin{cases} \begin{bmatrix} \dot{\hat{x}}_1 \\ \dot{\hat{x}}_2 \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} u - \begin{bmatrix} A_{12} \\ A_{22}^{-s} \end{bmatrix} e_y + \begin{bmatrix} 0 \\ \nu \end{bmatrix} \\ \hat{y}_i = \hat{x}_2 + c \\ \nu \triangleq -M\mathrm{sgn}(e_y) \end{cases} \tag{5}$$

where $A_{22}^{-s} = A_{22} - A_{22}^{s}$, $A_{22}^{s} \prec 0$ is the linear gain, the diagonal matrix $M \succ 0$ is the switching gain, and $e_y \triangleq \hat{y}_i - y_i$.

Based on the system dynamics of equation (4) and the SMO of equation (5), the observer error dynamics become

$$\begin{bmatrix} \dot{e}_1 \\ \dot{e}_2 \end{bmatrix} = \begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22}^{s} \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \end{bmatrix} + \begin{bmatrix} A_{12} \\ A_{22}^{-s} \end{bmatrix} \zeta_i - \begin{bmatrix} E_1 \\ E_2 \end{bmatrix} \eta - \begin{bmatrix} F_1 \\ F_2 \end{bmatrix} \Delta u_{i-1} + \begin{bmatrix} 0 \\ \nu \end{bmatrix} . \tag{6}$$

where $e_1 \triangleq \hat{x}_1 - x_1$ and $e_2 \triangleq \hat{x}_2 - x_2$. Furthermore, e_y can be expressed in terms of e_2 as $e_y = e_2 - \zeta_i$.

Lemma 1 presents bounds on e_1, e_2, and $\dot{e}_2$ in healthy and attacked conditions. These bounds are used in the detection logic design and detectability analysis in section IV.

Lemma 1: Define $\underline{e}_1 \leq e_1 \leq \bar{e}_1$, $\max(|\underline{e}_1|, |\bar{e}_1|) = \tilde{e}_1$, $|e_2| \leq \tilde{e}_2$, and $\underline{\dot{e}}_2 \leq |\dot{e}_2| \leq \bar{\dot{e}}_2$. Furthermore, denote bounds in healthy conditions, when Δu_{i-1}=0, with a superscript 0.

If, elementwise, $\mathrm{diag}(M) > |A_{21}|\tilde{e}_1 + |A_{22}|\bar{\zeta}_i + |E_2|\bar{\eta} + |F_2|\bar{\Delta}$, then the following conditions hold

1) $\bar{e}_1 = \bar{e}_1^0 - r_\Delta(\Delta u_{i-1})$
2) $\underline{e}_1 = \underline{e}_1^0 - r_\Delta(\Delta u_{i-1})$
3) $\bar{e}_1^0 = e^{A_{11}t}e_1(0) - A_{11}^{-1}(I - e^{A_{11}t})(|A_{12}|\bar{\zeta}_i + |E_1|\bar{\eta})$
4) $\underline{e}_1^0 = e^{A_{11}t}e_1(0) + A_{11}^{-1}(I - e^{A_{11}t})(|A_{12}|\bar{\zeta}_i + |E_1|\bar{\eta})$
5) $\tilde{e}_2 = \bar{\zeta}_i$
6) $\bar{\dot{e}}_2^0 = |A_{21}|\bar{e}_1^0 + |A_{22}^{-s}|\bar{\zeta}_i + |E_2|\bar{\eta} + |A_{22}^{s}|\bar{e}_2 + M$
7) $\underline{\dot{e}}_2^0 = |A_{21}|\underline{e}_1^0 - |A_{22}^{-s}|\bar{\zeta}_i - |E_2|\bar{\eta} - |A_{22}^{s}|\bar{e}_2 + M$

8) $\bar{\dot{e}}_2 = \bar{\dot{e}}_2^0 + \mathrm{sgn}(e_y) r(\Delta u_{i-1})$
9) $\underline{\dot{e}}_2 = \underline{\dot{e}}_2^0 + \mathrm{sgn}(e_y) r(\Delta u_{i-1})$
10) $\mathrm{sgn}(\dot{e}_2) = -\mathrm{sgn}(e_y)$

where $r_\Delta(\Delta u_{i-1}) = \int_0^t F_1 \Delta u_{i-1}(s) e^{A_{11}(t-s)} \mathrm{d}s$ and $r(\Delta u_{i-1}) = A_{21} r_\Delta(\Delta u_{i-1}) + F_2 \Delta u_{i-1}$.

Proof: Statements 3)-7) and 10) are proven [14]. Statements 1), 2), 8) and 9) are proven in the appendix. ■

IV. Detection Logic Design

In this section, the novel cyber-attack detection method will be described. This method directly analyses the behaviour of observer error e_2, and uses this to detect cyber-attacks. For comparison, the EOI based detection method presented in previous work [14] is presented in subsection IV-C. The detection performance of the two methods in a collaborative IC scenario will be compared in section V.

The novel proposed detection logic uses thresholds on the observer error e_2 based on the bounds in Lemma 1, and the relation $e_y = e_2 - \zeta_i$. The resulting thresholds, $\underline{e}_2 \leq e_2 \leq \bar{e}_2$, will be used for cyber-attack detection. Preferably one would directly monitor this condition, and detect a cyber-attack when it is violated. However, as e_2 is not known to the observer, this is not possible. Alternatively, the condition

$$\underline{e}_2 > \bar{e}_2 \tag{7}$$

can be monitored. Satisfying this condition implies violation of $\underline{e}_2 \leq e_2 \leq \bar{e}_2$, and can thus serve as detection condition.

A. Design of Error-bounds

At all time, $|e_2| < \bar{\zeta}_i$ (Lemma 1.5), and when a new measurement arrives to the observer $e_y - \bar{\zeta}_i \leq e_2 \leq e_y + \bar{\zeta}_i$. Denote the sequence of measurement times as $\{t_m\}$, which do not need to be equidistant. Then, for any time t_m

$$\begin{aligned} \bar{e}_2(t_m) &= \min(e_y(t_m) + \bar{\zeta}_i, \bar{\zeta}_i) \\ \underline{e}_2(t_m) &= \max(e_y(t_m) - \bar{\zeta}_i, -\bar{\zeta}_i) \end{aligned} \tag{8}$$

Furthermore, bounds on $\dot{e}_2^0$ are known from Lemma 1. With these, e_2 can be bound during each period $[t_{m-1}\ t_m]$ as

$$\begin{aligned} &\text{If } e_y(t_{m-1}) < 0 \\ &\quad \bar{e}_2(t) = \int_{t_{m-1}}^{t} \bar{\dot{e}}_2^0(T)\mathrm{d}T; \underline{e}_2(t) = \int_{t_{m-1}}^{t} \underline{\dot{e}}_2^0(T)\mathrm{d}T\,. \\ &\text{If } e_y(t_{m-1}) > 0 \\ &\quad \bar{e}_2(t) = -\int_{t_{m-1}}^{t} \underline{\dot{e}}_2^0(T)\mathrm{d}T; \underline{e}_2(t) = -\int_{t_{m-1}}^{t} \bar{\dot{e}}_2^0(T)\mathrm{d}T\,. \end{aligned} \tag{9}$$

The above bounds require further inspection. At first sight they seem to depend only on the modeled healthy system behaviour through $\bar{\dot{e}}_2^0$ and $\underline{\dot{e}}_2^0$. However, the bounds also depend on the real behaviour through e_y. The integration duration is dictated by the sign of e_y.The two bounds in equations (8) and (9) can be combined for $m \geq 1$ as

$$\begin{aligned} &\text{If } e_y(t_{m-1}) < 0 \\ &\quad \bar{e}_2(t_m) = \min(\int_{t_{m-1}}^{t_m} \bar{\dot{e}}_2^0(T)\mathrm{d}T, e_y(t_m) + \bar{\zeta}_i, \bar{\zeta}_i) \\ &\quad \underline{e}_2(t_m) = \max(\int_{t_{m-1}}^{t_m} \underline{\dot{e}}_2^0(T)\mathrm{d}T, e_y(t_m) - \bar{\zeta}_i, -\bar{\zeta}_i) \end{aligned}$$

$$\begin{aligned} &\text{If } e_y(t_{m-1}) > 0 \\ &\quad \bar{e}_2(t_m) = \min(-\int_{t_{m-1}}^{t_m} \underline{\dot{e}}_2^0(T)\mathrm{d}T, e_y(t_m) + \bar{\zeta}_i, \bar{\zeta}_i) \\ &\quad \underline{e}_2(t_m) = \max(-\int_{t_{m-1}}^{t_m} \bar{\dot{e}}_2^0(T)\mathrm{d}T, e_y(t_m) - \bar{\zeta}_i, -\bar{\zeta}_i) \end{aligned}$$

Equation (8) can be used to obtain $\bar{e}_2(t_0)$ and $\underline{e}_2(t_0)$. Based on these bounds, the detection criterion (7) can be monitored at every measurement time t_m.

B. Detectability Analysis

In this section, conditions are presented for which the proposed novel detection method can detect an attack. Furthermore, it is proven that in healthy conditions, the approach will never cause a detection.

First, introduce an assumption which is required in the presented proofs. This assumption is a relaxation of the matching condition commonly used in SMO literature.[14]

Assumption 5: $(F_2 - A_{21}A_{11}^{\dagger}F_1)$ is full column rank.

First, it will be proven that no detection occurs in healthy conditions.

Theorem 1: Consider system (4), observer (5) and detection criterion (7). In healthy conditions, i.e. if $\Delta u_{i-1} = 0\ \forall t$, the detection criterion will never be satisfied.

Proof: Define the sequence $\{t_{s_i}\}$ as the times where e_y changes sign, $\dot{e}_2^+$ as the average $|\dot{e}_2|$ while $e_y > 0$, and $\dot{e}_2^-$ as the average $|\dot{e}_2|$ while $e_y < 0$. Furthermore, without loss of generality, assume e_y becomes positive at every $t_{s_{2i}}$ allowing to write $t_i^+ = t_{s_{2i+1}} - t_{s_{2i}}$ and $t_i^- = t_{s_{2i+2}} - t_{s_{2i+1}}$.

Then, denote for the true dynamics of e_2 as

$$\begin{aligned} e_2(t_{s_{2i+2}}) &= e_2(t_{s_{2i}}) + c_i\,, \\ e_2(t_{s_{2i+2N}}) &= e_2(t_{s_{2i}}) + \sum_{j=0}^{N} c_{i+j}\ \forall N \in \mathbb{Z}\,, \end{aligned} \tag{10}$$

where $c_i = t_i^- \dot{e}_2^- - t_i^+ \dot{e}_2^+$. Now c_i can be bounded, using the bounds on e_2 from lemma 1 and $e_y = e_2 - \zeta_i$, as

$$\begin{aligned} -e_2(t_{s_{2i}}) + \max(e_y(t_{s_{2i}}) - \bar{\zeta}_i, -\bar{\zeta}_i) &\leq \sum_{j=i}^{N} c_i \\ &\leq -e_2(t_{s_{2i}}) + \min(e_y(t_{s_{2i}}) + \bar{\zeta}_i, +\bar{\zeta}_i)\ \forall N \in \mathbb{Z}\,. \end{aligned}$$

Furthermore, from equation (10), it can be derived that $\frac{t_i^+}{t_i^-} = \frac{\dot{e}_2^-}{\dot{e}_2^+} + \frac{c_i}{t_i^- \dot{e}_2^+}$. With this, $\bar{e}_2$ in equation (9) can be rewritten as

$$\bar{e}_2(t_{s_{2i+2}}) = \bar{e}_2(t_{s_{2i}}) + \frac{t_i^-}{\dot{e}_2^+}(\bar{\dot{e}}_2^0 \dot{e}_2^- - \underline{\dot{e}}_2^0 \dot{e}_2^+) + \frac{\bar{\dot{e}}_2^0}{\dot{e}_2^+} c_i\,.$$

which can be extended for $\bar{e}_2(t_{s_{2i+2N}})$ as

$$\bar{e}_2(t_{s_{2i+2N}}) = \bar{e}_2(t_{s_{2i}}) + \textstyle\sum_{j=0}^{N-1} \left(\frac{t_{i+j}^-}{\dot{e}_2^+}(\bar{\dot{e}}_2^0 \dot{e}_2^- - \underline{\dot{e}}_2^0 \dot{e}_2^+) + \frac{\bar{\dot{e}}_2^0}{\dot{e}_2^+} c_{i+j} \right), \tag{11}$$

for any $N \in \mathbb{Z}$. Similarly for $\underline{e}_2(t_{s_{2i+2N}})$ we can derive

$$\underline{e}_2(t_{s_{2i+2N}}) = \underline{e}_2(t_{s_{2i}}) + \textstyle\sum_{j=0}^{N-1} \left(\frac{t_{i+j}^-}{\dot{e}_2^+}(\underline{\dot{e}}_2^0 \dot{e}_2^- - \bar{\dot{e}}_2^0 \dot{e}_2^+) + \frac{\underline{\dot{e}}_2^0}{\dot{e}_2^+} c_{i+j} \right). \tag{12}$$

It can be seen that in healthy conditions, when $\underline{\dot{e}}_2^0 \leq \dot{e}_2^- \leq \bar{\dot{e}}_2^0$

and $\underline{\dot{e}}_2^0 \leq \dot{e}_2^+ \leq \dot{\bar{e}}_2^0$,

$$\begin{aligned}\bar{e}_2(t_{s_{2i+2N}}) - \bar{e}_2(t_{s_{2i}}) &\geq \sum_{j=0}^{N-1} \frac{\dot{\bar{e}}_2^0}{\dot{e}_2^+} c_{i+j} \geq \sum_{j=0}^{N-1} c_{i+j} \\ \underline{e}_2(t_{s_{2i+2N}}) - \underline{e}_2(t_{s_{2i}}) &\leq \sum_{j=0}^{N-1} \frac{\underline{\dot{e}}_2^0}{\dot{e}_2^+} c_{i+j} \leq \sum_{j=0}^{N-1} c_{i+j}\end{aligned} \tag{13}$$

By subtracting these inequalities it can be found that $\bar{e}_2(t_{s_{2i+2N}}) - \underline{e}_2(t_{s_{2i+2N}}) \geq \bar{e}_2(t_{s_{2i}}) - \underline{e}_2(t_{s_{2i}})$, i.e. considering the behaviour in equation (9), the difference between $\bar{e}_2$ and $\underline{e}_2$ is non-decreasing. This only leaves to prove that no detection occurs if the bounds are affected by equation (8).

If both bounds are affected by equation (8), $\bar{e}_2 - \underline{e}_2 = 2\bar{\zeta}_i - |e_y| \geq 0$. If only the lower bound is affected, use equation (13) to derive

$$\begin{aligned}\bar{e}_2(t_{s_{2i+2N}}) &\geq \bar{e}_2(t_{s_{2i}}) - e_2(t_{s_i}) + \max(e_y(t_{s_{2i}}) - \bar{\zeta}_i, -\bar{\zeta}_i) \\ &\geq \max(e_y(t_{s_{2i}}) - \bar{\zeta}_i, -\bar{\zeta}_i) \geq \underline{e}_2(t_{s_{2i+2N}})\end{aligned}$$

This proves the theorem. ■

Then two lemmas are introduced to support the proof of theorem 2, where sufficient conditions for attack detection are presented.

Lemma 2: consider $r(\Delta u_{i-1})$ as defined in Lemma 1. Then the following statements can be proven

1) $r(\Delta u_{i-1}) = 0 \; \forall t$ if $\Delta u_{i-1} = 0 \; \forall t$, i.e. healthy conditions.
2) There always exists $\gamma > 0$ such that within finite time $|r(\Delta u_{i-1}) - (F_2 - A_{21}A_{11}^{\dagger}F_1)\Delta u_{i-1}| \leq \gamma$.

Proof: By substituting $\Delta u_{i-1} = 0 \; \forall t$ in the function for $r(\Delta u_{i-1})$, it can directly be seen that $r(\Delta u_{i-1}) = 0 \; \forall t$. This proves statement a). For a constant Δu_{i-1},

$$\lim_{t\to\infty} r(\Delta u_{i-1}) = (F_2 - A_{21}A_{11}^{\dagger}F_1)\Delta u_{i-1}\,.$$

As $r(\Delta u_{i-1})$ is a smooth function, this means that within finite time $|r(\Delta u_{i-1}) - (F_2 - A_{21}A_{11}^{\dagger}F_1)\Delta u_{i-1}| < \gamma$ ■

Lemma 3: Consider the behaviours of $\bar{e}_2$ and $\underline{e}_2$ from equation (11) and (12). Assume there exist $\varepsilon^+ > 0$ and $\varepsilon^- > 0$ such that

- $\dot{\bar{e}}_2^0 < \dot{e}_2^+ - \varepsilon^-$ and $\underline{\dot{e}}_2^0 > \dot{e}_2^- + \varepsilon^-$ for the period $\left[t_{s_{2i}} \; t_{s_{2i+2N}}\right]$.
- OR $\underline{\dot{e}}_2^0 > \dot{e}_2^+ + \varepsilon^+$ and $\dot{\bar{e}}_2^0 < \dot{e}_2^- - \varepsilon^+$ for the period $\left[t_{s_{2i}} \; t_{s_{2i+2N}}\right]$.

Then, there exists an ε such that $\bar{e}_2(t_{s_{2i+2N}}) < \underline{e}_2(t_{s_{2i+2N}})$, if $N > \frac{4\bar{\zeta}_i}{\phi\varepsilon}$, where $\phi \leq \frac{t_{i+j}^-}{\dot{e}_2^+} \; \forall i,j$.

Proof: First, use $\dot{\bar{e}}_2^0 < \dot{e}_2^+ - \varepsilon^-$ and $\underline{\dot{e}}_2^0 > \dot{e}_2^- + \varepsilon^-$ to derive

$$\dot{\bar{e}}_2^0\dot{e}_2^- - \underline{\dot{e}}_2^0\dot{e}_2^+ < -(\dot{e}_2^- + \dot{e}_2^+)\varepsilon^- < -\varepsilon\,.$$

Then substitute $\dot{\bar{e}}_2^0\dot{e}_2^- - \underline{\dot{e}}_2^0\dot{e}_2^+ < -\varepsilon$ and $\phi \leq \frac{t_{i+1}^-}{\dot{e}_2^+} \; \forall i,j$ in equation (11) giving

$$\bar{e}_2(t_{s_{2i+2N}}) - \bar{e}_2(t_{s_{2i}}) < -N\varepsilon\phi + \sum_{j=0}^{N-1} c_{i+j}\,.$$

Using the bound on c_{i+j} gives

$$\begin{aligned}\bar{e}_2(t_{s_{2i+2N}}) &< \bar{e}_2(t_{s_{2i}}) - N\varepsilon\phi - e_2(t_{s_{2i}}) + \min(e_y(t_{s_{2i}}) + \bar{\zeta}_i, \bar{\zeta}_i) \\ &< -N\varepsilon\phi + 2\min(e_y(t_{s_{2i}}) + \bar{\zeta}_i, \bar{\zeta}_i) - \max(e_y(t_{s_{2i}}) - \bar{\zeta}_i, -\bar{\zeta}_i)\end{aligned}$$

Meanwhile, always $\underline{e}_2(t_{s_{2i+2N}}) > \max(e_y(t_{s_{2i}}) - \bar{\zeta}_i, -\bar{\zeta}_i)$, which with some simplification leads to

$$\bar{e}_2(t_{s_{2i+2N}}) - \underline{e}_2(t_{s_{2i+2N}}) < -N\varepsilon\phi + 4\bar{\zeta}_i$$

So, $\bar{e}_2(t_{s_{2i+2N}}) < \underline{e}_2(t_{s_{2i+2N}})$ if $N > \frac{4\bar{\zeta}_i}{\varepsilon\phi}$.

The same result can be obtained by using $\underline{\dot{e}}_2^0 > \dot{e}_2^+ + \varepsilon^+$ and $\dot{\bar{e}}_2^0 < \dot{e}_2^- - \varepsilon^+$ to obtain $\underline{\dot{e}}_2^0\dot{e}_2^- - \dot{\bar{e}}_2^0\dot{e}_2^+ > \varepsilon$ and substituting in equation (12). ■

Theorem 2: Consider system (4), with observer (5) and detection criterion (7). There exist a δ and τ such that the detection condition (7) will be satisfied if $|r(\Delta u_{i-1})| \geq \delta$ for at least a duration τ. Furthermore, if assumption 5 holds, there always exists a Δu_{i-1} such that $|r(\Delta u_{i-1})| \geq \delta$.

Proof: In Lemma 3 conditions on $\dot{e}_2^+$ and $\dot{e}_2^-$ are presented such that detection occurs within a duration $\tau = t_{s_{2i+2N}} - t_{s_{2i}}$. Here N is defined in Lemma 3. In this proof it thus remains to be shown that there exists a δ such that the conditions on $\dot{e}_2^+$ and $\dot{e}_2^-$ from Lemma 3 hold for any attack $r(\Delta u_{i-1}) \geq \delta$.

From equation (14), use $\dot{e}_2^+ = \dot{e}_2^0 + r(\Delta u_{i-1})$ and $\dot{e}_2^- = \dot{e}_2^0 - r(\Delta u_{i-1})$. If $r(\Delta u_{i-1}) > \dot{\bar{e}}_2^0 - \underline{\dot{e}}_2^0 + \varepsilon^-$, then $\dot{e}_2^+ > \dot{e}_2^0 + \dot{\bar{e}}_2^0 - \underline{\dot{e}}_2^0 + \varepsilon^- \geq \dot{\bar{e}}_2^0 + \varepsilon^-$ and $\dot{e}_2^- < \dot{e}_2^0 - \dot{\bar{e}}_2^0 + \underline{\dot{e}}_2^0 - \varepsilon^- \leq \underline{\dot{e}}_2^0 - \varepsilon^-$. This is equivalent to the first condition in Lemma 3. The second condition holds if $r(\Delta u_{i-1}) < \underline{\dot{e}}_2^0 - \dot{\bar{e}}_2^0 - \varepsilon^+$ and can be proven similarly.

Furthermore, in Lemma 2 it is shown that there exists a $\gamma > 0$ such that $|r(\Delta u_{i-1}) - (F_2 - A_{21}A_{11}^{\dagger}F_1)\Delta u_{i-1}| \leq \gamma$ within finite time. Therefore, if assumption 5 holds, there always exists a Δu_{i-1} to obtain $|r(\Delta u_{i-1})| \geq \delta$. ■

C. Equivalent Output Injection based detection

In this subsection the equivalent output injection (EOI) based detection, as previously introduced in [14], is presented for comparison with the novel detection method. First the EOI is defined as

$$\dot{\nu}_{\text{fil}} = K(\nu - \nu_{\text{fil}})\,,$$

where ν_{fil} is the EOI, and $K \succ 0$ is a diagonal gain matrix.

The EOI was originally introduced to estimate the cyber-attack. In [14] the following was proven.

Proposition 1: Consider noiseless system (4), where $\zeta_i = 0$, and the SMO (5). If $\text{diag}(M) > |A_{21}|\tilde{e}_1 + |A_{22}|\bar{\zeta}_i + |E_2|\bar{\eta} + |F_2|\bar{\Delta}$, then $e_2 \to 0$ and $\dot{e}_2 \to 0$. Furthermore, assuming a constant cyber-attack, the cyber-attack estimate

$$\hat{\Delta}u_{i-1} = (F_2 - A_{21}A_{11}^{-1}F_1)^{\dagger}\nu_{\text{fil}}$$

has an accuracy of

$$|\Delta u_{i-1} - \hat{\Delta}u_{i-1}| \leq |(F_2 - A_{21}A_{11}^{-1}F_1)^{\dagger}(A_{21}A_{11}^{-1}|E_1| + |E_2|)\bar{\eta}|$$

■

The EOI can also be used for cyber-attack detection. Based on the bounds presented in Lemma 1, a threshold for EOI-based cyber-attack detection is introduced in [14], globally bounding the healthy EOI behaviour. In the threshold, each element $_{(i)}$ is defined as

$$\bar{\nu}_{\text{fil},(i)} = e^{-k\bar{t}_{(i)}^{0,*}}\bar{U}_{(i)} + (1 - e^{-k\bar{t}_{(i)}^{0,*}})m\,,$$

where $k = K_{(i,i)}$, $m = M_{(i,i)}$, $\bar{t}^{0,*} = \lim_{t\to\infty} \frac{2\bar{e}_2}{\underline{\dot{e}}_2^0}$, and $\bar{U} = \lim_{t\to\infty} |A_{12}|\bar{e}_1^0 + |A_{22}^{-s}|\bar{\zeta}_i + |E_2|\bar{\eta} + |A_{22}^s|\bar{e}_2$. A lower threshold

$\underline{v}_{\text{fil}} = -\bar{v}_{\text{fil}}$ is derived similarly. A cyber-attack is detected if the condition $\underline{v}_{\text{fil}} \leq v_{\text{fil}} \leq \bar{v}_{\text{fil}}$ is violated.

V. Simulation of Intersection Control

A simulation is performed with 2 cars approaching an intersection. The car closest to the intersection will be referred to as the leader car, for which the input sequence is pre-defined. The car furthest from the intersection is the follower car, and is controlled using the control law from [24] shown below. The detection algorithm works regardless of the control law.

$$\dot{u}_i = -\frac{1}{h}(u_i + k_p \varepsilon_1 + k_d \dot{\varepsilon}_1 - u_{i-1}).$$

Here $\varepsilon_1 = d_i - d_{i,r}$, $d_i = p_i - p_{i-1} - L_{i-1}$ is the relative distance of the cars to the intersection, and $d_{i,r} = r + hv_i$ is the desired relative distance. The time headway h, standstill distance r, and control gains k_p and k_d, are defined in Table I.

TABLE I

Parameters used in simulation

Parameter	Value	Parameter	Value
$p_0(0)$	$-40\,[m]$	$p_1(0)$	$-50\,[m]$
$v_0(0)$	$8\,\left[\frac{m}{s}\right]$	$v_1(0)$	$10\,\left[\frac{m}{s}\right]$
$a_0(0)$	$0\,\left[\frac{m}{s^2}\right]$	$a_1(0)$	$0\,\left[\frac{m}{s^2}\right]$
τ_0	$0.11\,[s]$	τ_1	$0.1\,[s]$
L_1	$4\,[m]$	r_τ	$0.9\,[-]$
h	$0.7\,[s]$	r	$1.5\,[m]$
k_p	$0.2\,[s^{-2}]$	k_d	$0.7\,[s^{-1}]$
$\bar{\Delta}$	$10\,\left[\frac{m}{s^2}\right]$	$\bar{\eta}$	$1\,\left[\frac{m}{s^2}\right]$
$\bar{\zeta}_1$	$[0.15\ 0.3\ 0.03\ 0.15]^\top\,\left[m\ \frac{m}{s}\ \frac{m}{s}\ \frac{m}{s^2}\right]$		
K	$I_4\,[s^{-1}]$	A^s_{22}	$-0.1 \cdot I_4\,[s^{-1}]$
M	$\text{diag}([0.5\ 11.5\ 0.2\ 2.0])\,\left[m\ \frac{m}{s}\ \frac{m}{s}\ \frac{m}{s^2}\right]$		

Based on [25], IC is initiated when the cars are within $50[m]$ from the intersection. Furthermore, the intersection is approached at $8[m/s] \approx 30[km/h]$, which is a common standard speed in urban areas. In figure 2, the input of the leader car, and the cyber-attack considered are show. It can be seen that the lead vehicle drives at a constant speed, and at $t = 0.5[s]$ a step attack is performed on the communication. In figure 3 it is shown that this attack causes the cars to gradually drive closer together, eventually leading to a crash at the intersection. The crash occurs when the follower vehicle enters the intersection, i.e. $p_f = -4\ [m]$. At this point the lead vehicle has not yet left the intersection, i.e. $-4\ [m] < p_l < 0\ [m]$, resulting in a crash. Figure 4 shows the detection results obtained with the novel detection logic presented in Section IV. Detection first occurs for a very short period at $t = 0.64\ [s]$. This is not visible in the figure. More consistent detection occurs at $t = 0.82\ [s]$. This consistent detection occurs well before the crash occurs at $t = 4.8\ [s]$.

In Figure 6, it is shown that the EOI based detection method from previous work also detects the attack. However, the detection only occurs at $t = 2.73\ [s]$. This is significantly slower than the detection with the novel detection method.

The capability of the EOI based method to also estimate the attack is illustrated in figure 5. This estimate of the attack can be very useful in designing an effective control strategy to deal with the attack.

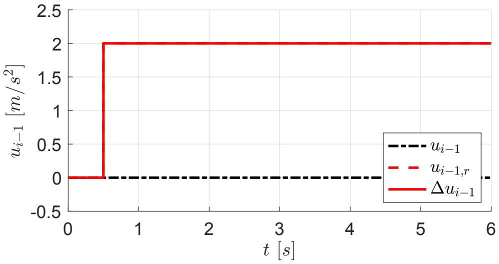

Fig. 2. Input of leader car and considered attack on the communication.

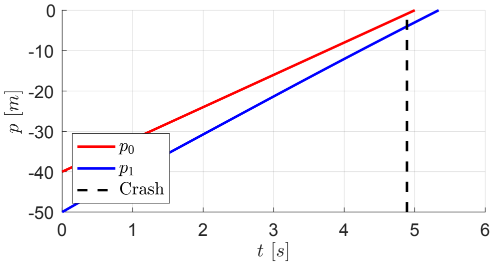

Fig. 3. Effect of the attack on vehicle positions, leading up to the crash.

As both detection methods depend on the same observer, it is feasible to implement both methods concurrently. In this way the best properties of both methods can be combined.

VI. Conclusions

Intersections, in the way they are currently designed for human drivers, are a major cause of accidents as well as traffic delays. Therefore, automated control of vehicles in intersections offers great potential for improvement. However, as Intersection Control systems do rely on wireless V2X networks for traffic coordination, security of such communication channel is paramount for safety. In V2X networks a first line of preventive security measures are already in place at the protocol level to make cyber-attacks more difficult. Still the possibility of an inside attacker, or one capable of overcoming preventive security measures cannot be ruled out. To protect against such a scenario, a second line of defenses based on a cyber-attack detection method is proposed in this paper. In particular, a novel detector based on a Sliding Mode Observer and a corresponding set of thresholds was designed. With respect to previous results, the novel detection approach is shown to be faster and more sensitive, as the filtering of the observer Equivalent Output Injection is avoided. Theoretical results certifying the robustness and detectability of the proposed approach were provided, as well as a simulation study.

In the future, a two-stage approach based on a fault detector and a fault identification scheme may be envisaged, thus paving the way for fully autonomous accommodation of faults and cyber-attacks in Intersection Control systems based on cooperative autonomous vehicles. Furthermore, adaptations to the detection method are envisioned for which boundedness of the attack is no longer required.

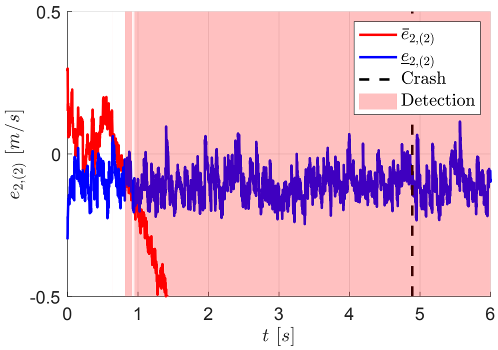

Fig. 4. Error bounds and detection performance of novel detection logic.

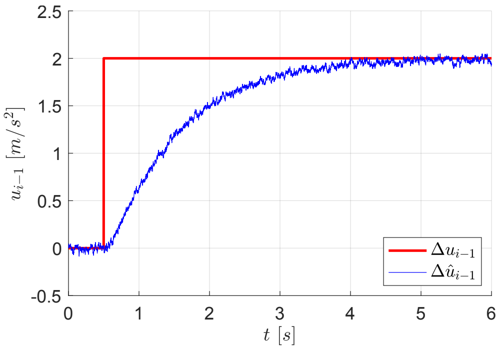

Fig. 5. Attack estimation by EOI based estimation method.

Appendix

From the first row of equation (6), using lemma 1.1.1 in [26] we obtain $e_1(t) = e^{A_{11}(t)}e_1(0) - r_\zeta(\zeta_i) - r_\eta(\eta) - r_\Delta(\Delta u_{i-1})$, where $r_\zeta(\zeta_i) = \int_0^t A_{21}\zeta_i(s)e^{A_{11}(t-s)}\mathrm{d}s$ and $r_\eta(\eta) = \int_0^t E_1\eta(s)e^{A_{11}(t-s)}\mathrm{d}s$.

From this it can be concluded that $e_1(t) = e_1^0(t) - r_\Delta(\Delta u_{i-1})$, and therefore $\bar{e}_1(t) = \bar{e}_1^0(t) - r_\Delta(\Delta u_{i-1})$ and $\underline{e}_1(t) = \underline{e}_1^0(t) - r_\Delta(\Delta u_{i-1})$, which proves statements 1) and 2) in Lemma 1.

Furthermore $\dot{e}_2$ from the second row of equation (6) is $\dot{e}_2 = A_{21}e_1 + A_{22}^s e_2 + A_{22}^{-s}\zeta_i - E_2\eta - F_2\Delta u_{i-1} - M\mathrm{sgn}(e_y)$. With this we can write,

$$\dot{e}_2 = \dot{e}_2^0 + A_{21}(e_1 - e_1^0) - F_2\Delta u_{i-1} = \dot{e}_2^0 - A_{21}r_\Delta(\Delta u_{i-1}) - F_2\Delta u_{i-1}. \quad (14)$$

To prove statements 8) and 9) in Lemma 1 we finally observe that $\bar{\dot{e}}_2 = \bar{\dot{e}}_2^0 + \mathrm{sgn}(e_y)r(\Delta u_{i-1})$ and $\underline{\dot{e}}_2 = \underline{\dot{e}}_2^0 + \mathrm{sgn}(e_y)r(\Delta u_{i-1})$, .

References

[1] European Commission, "Traffic safety basic facts on junctions," 2016.

[2] K. Dresner and P. Stone, "Multiagent traffic management: An improved intersection control mechanism," in *AAMAS'05, Procs of*, 2005.

[3] T. Arts, M. Dorigatti, and S. Tonetta, "Making Implicit Safety Requirements Explicit: An AUTOSAR Safety Case," in *Computer Safety, Reliability, and Security*, A. Bondavalli and F. Di Giandomenico, Eds. Cham: Springer International Publishing, 2014, vol. 8666, pp. 81–92.

[4] M. Amoozadeh, A. Raghuramu, C.-N. Chuah, D. Ghosal, H. M. Zhang, J. Rowe, and K. Levitt, "Security vulnerabilities of connected vehicle streams and their impact on cooperative driving," *IEEE Communications Magazine*, vol. 53, no. 6, pp. 126–132, 2015.

[5] C. Yan, W. Xu, and J. Liu, "Can you trust autonomous vehicles: Contactless attacks against sensors of self-driving vehicle," *DEF CON*, vol. 24, no. 8, p. 109, 2016.

[6] R. G. Engoulou, M. Bellaïche, S. Pierre, and A. Quintero, "Vanet security surveys," *Computer Communications*, vol. 44, pp. 1–13, 2014.

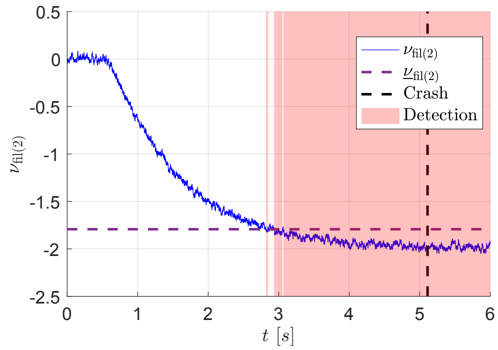

Fig. 6. Attack detection by EOI based detection method.

[7] R. van der Heijden, T. Lukaseder, and F. Kargl, "Analyzing attacks on cooperative adaptive cruise control (cacc)," in *2017 IEEE Vehicular Networking Conference (VNC)*. IEEE, 2017, pp. 45–52.

[8] N. Bißmeyer, S. Mauthofer, K. M. Bayarou, and F. Kargl, "Assessment of node trustworthiness in vanets using data plausibility checks with particle filters," in *VNC, Procs. of*, 2012.

[9] S. Dietzel, J. Petit, G. Heijenk, and F. Kargl, "Graph-based metrics for insider attack detection in vanet multihop data dissemination protocols," *IEEE Trans. on Vehic. Tech.*, vol. 62, no. 4, pp. 1505–1518, 2012.

[10] G. Yan, S. Olariu, and M. C. Weigle, "Providing vanet security through active position detection," *Computer communications*, vol. 31, no. 12, pp. 2883–2897, 2008.

[11] F. Pasqualetti, F. Dörfler, and F. Bullo, "Attack detection and identification in cyber-physical systems," *IEEE Trans. on Automatic Control*, vol. 58, no. 11, pp. 2715–2729, 2013.

[12] Y. Quan, W. Chen, Z. Wu, and L. Peng, "Distributed fault detection and isolation for leader–follower multi-agent systems with disturbances using observer techniques," *Nonlinear Dynamics*, vol. 93, no. 2, pp. 863–871, 2018.

[13] N. Jahanshahi and R. M. Ferrari, "Attack detection and estimation in cooperative vehicles platoons: A sliding mode observer approach," in *NECSYS, Procs. of*, August 2018.

[14] T. Keijzer and R. M. Ferrari, "A sliding mode observer approach for attack detection and estimation in autonomous vehicle platoons using event triggered communication," in *CDC, Procs. of*, 2019.

[15] ——, "Detection of network and sensor cyber-attacks in platoons of cooperative autonomous vehicles: a sliding-mode observer approach," in *ECC, Procs. of*, 2021.

[16] A. I. M. Medina, N. van de Wouw, and H. Nijmeijer, "Cooperative intersection control based on virtual platooning," *IEEE Trans. on Intelligent Transp. Systems*, vol. 19, no. 6, pp. 1727–1740, 2017.

[17] G. R. Campos, P. Falcone, H. Wymeersch, R. Hult, and J. Sjöberg, "Cooperative receding horizon conflict resolution at traffic intersections," in *CDC, Procs. of*, 2014.

[18] M. A. S. Kamal, J.-i. Imura, T. Hayakawa, A. Ohata, and K. Aihara, "A vehicle-intersection coordination scheme for smooth flows of traffic without using traffic lights," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 3, pp. 1136–1147, 2014.

[19] A. Katriniok, P. Kleibaum, and M. Joševski, "Distributed model predictive control for intersection automation using a parallelized optimization approach," in *IFAC World Congress*, 2017.

[20] M. Kneissl, A. Molin, H. Esen, and S. Hirche, "A feasible mpc-based negotiation algorithm for automated intersection crossing," in *2018 European Control Conference (ECC)*. IEEE, 2018, pp. 1282–1288.

[21] Robert Bosch GmbH. (2021) Corner radar sensor. [Online]. Available: https://www.bosch-mobility-solutions.com/en/products-and-services/passenger-cars-and-light-commercial-vehicles/driver-assistance-systems/automatic-emergency-braking/corner-radar-sensor/

[22] E. Janse, "Anomaly detection in intersection control: Sliding mode observer based anomaly detection in virtual platooning enabled intersection control," Master's thesis, TU Delft, 2020.

[23] C. Edwards, S. K. Spurgeon, and R. J. Patton, "Sliding mode observers for fault detection and isolation," *Automatica*, vol. 36, no. 4, pp. 541–553, Apr. 2000.

[24] 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 *14th International IEEE Conference on Intelligent Transportation Systems*, 2011, pp. 260–265.

[25] M. Abdulla, E. Steinmetz, and H. Wymeersch, "Vehicle-to-vehicle communications with urban intersection path loss models," in *2016 IEEE Globecom Workshops (GC Wkshps)*, 2016, pp. 1–6.

[26] V. Lakshmikantham, S. Leela, and A. A. Martynyuk, *Stability Analysis of Nonlinear Systems*. Birkhäuser, 2015.