A Self-Triggered Control Watermarking Scheme for Detecting Replay Attacks

Wolleswinkel, Bart; Ferrari, Riccardo; Mazo, M.

DOI
10.1109/CDC56724.2024.10886758

Publication date
2025

Document Version
Final published version

Published in
Proceedings of the IEEE 63rd Conference on Decision and Control, CDC 2024

Citation (APA)
Wolleswinkel, B., Ferrari, R., & Mazo, M. (2025). A Self-Triggered Control Watermarking Scheme for Detecting Replay Attacks. In *Proceedings of the IEEE 63rd Conference on Decision and Control, CDC 2024* (pp. 4568-4573). (Proceedings of the IEEE Conference on Decision and Control). IEEE. https://doi.org/10.1109/CDC56724.2024.10886758

2024 IEEE 63rd Conference on Decision and Control (CDC)
December 16-19, 2024. MiCo, Milan, Italy

A Self-Triggered Control Watermarking Scheme for Detecting Replay Attacks

Bart Wolleswinkel, Riccardo Ferrari, and Manuel Mazo Jr.

2024 IEEE 63rd Conference on Decision and Control (CDC) | 979-8-3503-1633-9/24/$31.00 ©2024 IEEE | DOI: 10.1109/CDC56724.2024.10886758

Abstract— **We propose a novel watermarking scheme by modifying a self-triggered control (STC) policy, aimed at detecting replay attacks for linear time-invariant (LTI) systems. We show that by employing non-deterministic early triggering of the STC mechanism, replay attacks can be detected by a modified χ^2 detector which takes into account the aperiodic nature of the inter-sample times. Specifically, we consider the case where a periodic reference signal is tracked, which makes these systems vulnerable to replay attacks. The proposed approach is modular and can be retrofitted to legacy systems. An approach for designing an online optimal early triggering mechanism is provided. This is validated through an illustrative numerical example in which we compare our method to scenarios employing both additive and multiplicative watermarking.**

I. Introduction

In the last decades, advances in communication technology have significantly impacted industrial control systems (ICSs), such as manufacturing, energy generation, and water sanitation. Control systems in which the physical process is tightly interconnected with digital processes have been named cyber-physical systems (CPSs). Increasingly often, CPSs make use of a (wireless) communication network between controller and plant, so-called networked control systems (NCSs), in part due to the flexibility they offer.

Despite such advantages, NCSs also pose distinct challenges. For one, the bandwidth of the underlying communications network is often severely limited. As such, aperiodic control techniques have been developed. In particular, event-triggered control (ETC) [1] and self-triggered control (STC) [2] have attracted much attention. Aperiodic feedback does entail that each communication becomes more critical: simultaneously, this aperiodicity provides an additional degree of freedom, which can, for instance, be exploited for the detection of anomalies.

Alongside the aforementioned communication challenges, CPSs are also exposed to new risks. Digitalization allows an adversary to alter the digital control signals, the effects of which can propagate to the physical process, possibly causing physical damage, severe economic disruption, and even loss of human life [3]. These attacks on CPSs distinguish themselves from faults in that they are crafted with intent and are adversarial in nature [4], [5]. In this context, research has addressed both the design of diverse types of attacks [6] as well as countermeasures [7], laying the foundations for the field of *secure control*.

A particular class of attacks against which, by definition, classical anomaly detection methods are ineffective are so-called *stealthy attacks* [8], [9]. Of these, replay attacks have received significant attention in the literature [10], [11], in part due to the fact that an adversary does not require detailed knowledge of the plant model [12], and due to their real-world precedent [13]. To counter the limitation of passive detection methods, *active* diagnosis techniques have been developed. Examples include output measurement coding [14], additive watermarking [15], moving target defense [16], and multiplicative watermarking [17]. Whilst effective, additive approaches suffer from a degradation of the control performance, while multiplicative ones remedy this problem at the cost of requiring additional computation power at the sensor side [17].

The literature on secure control of ETC and STC systems is not yet mature. Whilst denial-of-service (DoS) attacks have received considerable attention in the ETC literature [18], [19], other types of attacks are still an area ripe for novel development. Replay attacks on aperiodic sampled systems have been, for instance, considered in [20], [21]. However, [21] only considers the noise-free case, whilst [20] relies on a shared secret between sensors and controllers.

In this manuscript, we explore a different paradigm by focusing on stochastic systems. In particular, we propose a computationally light watermarking scheme that does not rely on a shared secret and can be implemented using general quadratic triggering conditions. To the best of the authors' knowledge, this is the first time a watermarking strategy based on explicitly manipulating the inter-sample times of an STC mechanism is proposed.

Notation: The operator $\operatorname{col}(\bullet)$ concatenates its operands vertically such that $\operatorname{col}(\boldsymbol{v}_1, \boldsymbol{v}_2) = [\ \boldsymbol{v}_1^{\mathrm{T}} \quad \boldsymbol{v}_2^{\mathrm{T}}\ ]^{\mathrm{T}}$. Given a set $\mathbb{A}$, we define $c \cdot \mathbb{A} = \{c \cdot a \,|\, a \in \mathbb{A}\}$. For a symmetric matrix $\boldsymbol{W} = \boldsymbol{W}^{\mathrm{T}}$, let $\boldsymbol{W} \succ 0$ ($\boldsymbol{W} \succcurlyeq 0$) denote that $\boldsymbol{W}$ is positive definite (positive semidefinite), and let $\lambda_{\min}(\boldsymbol{W})$ denote its smallest eigenvalue. With $X_i \sim p_i$, we denote a random variable X_i with distribution p_i.

II. Problem definition

Let us consider the following discrete-time[1] linear time-invariant (LTI) plant:

$$\mathcal{P}: \quad \begin{aligned} \boldsymbol{x}[k+1] &= \boldsymbol{A}\boldsymbol{x}[k] + \boldsymbol{B}\boldsymbol{u}[k] + \boldsymbol{w}[k], && \text{(1a)} \\ \boldsymbol{y}[k] &= \boldsymbol{C}\boldsymbol{x}[k] + \boldsymbol{v}[k], && \text{(1b)} \end{aligned}$$

This work has been partially supported by the EU Horizon program through the project TWAIN, grant id 101122194.

The authors are with the Delft Center for Systems and Control (DCSC), Mechanical Engineering Faculty, TU Delft, Delft, The Netherlands (email:{b.wolleswinkel, r.ferrari, m.mazo}@tudelft.nl).

[1] For discretizing stochastic continuous-time systems, see e.g. [12].

with state vector $\boldsymbol{x}[n] \in \mathbb{R}^{n_\mathrm{x}}$, state matrix $\boldsymbol{A} \in \mathbb{R}^{n_\mathrm{x}\times n_\mathrm{x}}$, input matrix $\boldsymbol{B} \in \mathbb{R}^{n_\mathrm{x}\times n_\mathrm{u}}$, and output matrix $\boldsymbol{C} \in \mathbb{R}^{n_\mathrm{y}\times n_\mathrm{u}}$. The process noise $\boldsymbol{w}[k]$ and measurement noise $\boldsymbol{v}[k]$ are uncorrelated, i.i.d. Gaussian processes with zero mean and covariance matrices $\boldsymbol{\Sigma}_\mathrm{w} \succcurlyeq 0$ and $\boldsymbol{\Sigma}_\mathrm{v} \succcurlyeq 0$, respectively. We make the following standard assumptions:

Assumption 1. *The pair $(\boldsymbol{A}, \boldsymbol{B})$ is controllable and the pair $(\boldsymbol{A}, \boldsymbol{C})$ is observable.* ◊

The considered architecture can be seen in Fig. 1. The plant $\mathcal{P}$ and controller $\mathcal{C}$ are physically non-collocated, and transmission over the communications channel is based on an STC mechanism. Here, k_i denotes the i-th sample instant, with $i \in \mathbb{N}_0$. At those time instants, the sensors send a measurement $\boldsymbol{y}_i = \boldsymbol{y}[k_i]$ to the controller, which then computes an updated input $\boldsymbol{u}_\mathrm{c}[k_i]$ and transmits this to the actuators. Furthermore, we assume a hold mechanism at the plant side, resulting in

$$\boldsymbol{u}[k] = \boldsymbol{u}_\mathrm{c}[k_i], \quad \forall\, k_i \leqslant k < k_{i+i}. \tag{2}$$

The actuation signal $\boldsymbol{u}_\mathrm{c}$ is computed by a dynamic output-feedback controller $\mathcal{C}$, designed for the tracking of a reference signal $\boldsymbol{r}$, of the following form:

$$\mathcal{C}: \quad \begin{aligned} \boldsymbol{x}_\mathrm{c}[k+1] &= \boldsymbol{A}_\mathrm{c}\boldsymbol{x}_\mathrm{c}[k] + \boldsymbol{B}_\mathrm{c}\boldsymbol{e}[k], & \text{(3a)}\\ \boldsymbol{u}_\mathrm{c}[k] &= \boldsymbol{C}_\mathrm{c}\boldsymbol{c}[k] + \boldsymbol{D}_\mathrm{c}\boldsymbol{e}[k], & \text{(3b)} \end{aligned}$$

with $\boldsymbol{e}[k] := \boldsymbol{r}[k] - \boldsymbol{y}_i^{\downarrow}$ for $k_i \leqslant k < k_{i+1}$. Here, $\boldsymbol{y}_i^{\downarrow}$ denotes the received measurement whilst $\boldsymbol{y}_i$ denotes the transmitted measurement, which might be different due to the presence of an attack. We employ from here on the shorthand notation $\boldsymbol{u}_i = \boldsymbol{u}[k_i]$, and similarly for all other time signals.

Let $\kappa_{i+1} = k_{i+1} - k_i$ denote the time interval between the next and the i-th sample instant. Without loss of generality, we define $k_0, \kappa_0 := 0$. In the next section, we describe how the time instants k_i are determined. Hereinafter, we drop any dependence on k_i, and therefore κ_i, for brevity.

Finally, since we are considering output feedback, both the STC policy $\mathcal{S}$ and the detector $\mathcal{D}$ need an estimate of the state $\hat{\boldsymbol{x}}_i$. The dynamics at sampling instants, resulting from $\mathcal{P}$ with the aperiodic feedback (2), can be modeled as a switched linear (SL) system as follows:

$$\begin{aligned} \boldsymbol{x}_{i+1} &= \boldsymbol{A}_{\kappa_{i+1}}\boldsymbol{x}_i + \boldsymbol{B}_{\kappa_{i+1}}\boldsymbol{u}_i + \boldsymbol{w}_{i+1}, & \text{(4a)}\\ \boldsymbol{y}_i &= \boldsymbol{C}\boldsymbol{x}_i + \boldsymbol{v}[k_i], & \text{(4b)} \end{aligned}$$

with matrices $\boldsymbol{A}_\kappa$, $\boldsymbol{B}_\kappa$ given by

$$\boldsymbol{A}_\kappa = \boldsymbol{A}^\kappa, \qquad \boldsymbol{B}_\kappa = \sum_{\ell=0}^{\kappa-1} \boldsymbol{A}^\ell \boldsymbol{B}. \tag{5}$$

The resulting process noise signal $\boldsymbol{w}_i$ is again Gaussian with mean $\boldsymbol{0}$ and covariance matrix

$$\boldsymbol{\Sigma}_{\mathrm{w},\kappa} = \sum_{\ell=0}^{\kappa-1} \boldsymbol{A}^\ell \boldsymbol{\Sigma}_\mathrm{w} \left(\boldsymbol{A}^\mathrm{T}\right)^\ell. \tag{6}$$

Note that the random vectors $\boldsymbol{w}_i$ are no longer identically distributed for all i. However, they remain independent, i.e., $\mathbb{E}\left[\boldsymbol{w}_i^\mathrm{T}\boldsymbol{w}_{i'}\right] = \boldsymbol{0}$, $\forall i \neq i'$. We can therefore construct an estimator for the state at sampling instants, with a time-varying Kalman filter as follows:

$$\mathcal{O}: \quad \begin{aligned} \hat{\boldsymbol{x}}_{i\,|\,i-1} &= \boldsymbol{A}_{\kappa_i}\hat{\boldsymbol{x}}_{i-1} + \boldsymbol{B}_{\kappa_i}\boldsymbol{u}_{i-1}, & \text{(7a)}\\ \boldsymbol{\Sigma}_{\mathrm{x},i\,|\,i-1} &= \boldsymbol{A}_{\kappa_i}\boldsymbol{\Sigma}_{\mathrm{x},i-1}\boldsymbol{A}_{\kappa_i}^\mathrm{T} + \boldsymbol{\Sigma}_{\mathrm{w},\kappa_i}, & \text{(7b)}\\ \boldsymbol{\Sigma}_{\mathrm{z},i} &= \boldsymbol{C}\boldsymbol{\Sigma}_{\mathrm{x},i\,|\,i-1}\boldsymbol{C}^\mathrm{T} + \boldsymbol{\Sigma}_\mathrm{v}, & \text{(7c)}\\ \boldsymbol{H}_i &= \boldsymbol{\Sigma}_{\mathrm{x},i\,|\,i-1}\boldsymbol{A}_{\kappa_i}^\mathrm{T}\boldsymbol{\Sigma}_{\mathrm{z},i}^{-1}, & \text{(7d)}\\ \hat{\boldsymbol{x}}_i &= (\boldsymbol{I} - \boldsymbol{H}_i\boldsymbol{C})\hat{\boldsymbol{x}}_{i\,|\,i-1} + \boldsymbol{H}_i\boldsymbol{y}_i^{\downarrow}, & \text{(7e)}\\ \boldsymbol{\Sigma}_{\mathrm{x},i} &= (\boldsymbol{I} - \boldsymbol{H}_i\boldsymbol{A}_{\kappa_i})\boldsymbol{\Sigma}_{\mathrm{x},i\,|\,i-1}. & \text{(7f)} \end{aligned}$$

The full system architecture can be seen in Fig. 1. The design of the STC policy $\mathcal{S}$ and detector $\mathcal{D}$ are discussed in §III.

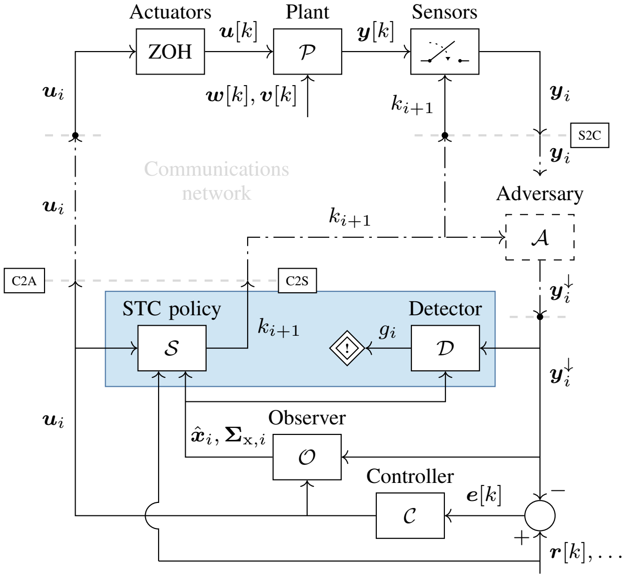

Fig. 1. The considered networked control system architecture. Novel contributions are highlighted in blue.

A. Adversary model

Let us define a safe region $\mathbb{X}_\mathrm{s} \subset \mathbb{R}^{n_\mathrm{x}}$, denoting the bounded set that ensures safety of the system. The adversary $\mathcal{A}$ is able to launch a *replay attack* [11], given by

$$\boldsymbol{y}_i^{\downarrow} = \begin{cases} \boldsymbol{y}_{i-\Delta i}, & i \geqslant I_\mathrm{a},\\ \boldsymbol{y}_i, & \text{otherwise}. \end{cases} \tag{8}$$

Here, I_a denotes the start of the attack, and $\Delta i \in \mathbb{N}$ denotes the amount of delay. A replay attack as in (8) can be successful if the control system exhibits some form of repetitiveness. Such is the case when the reference signal is periodic:

Assumption 2. *The reference signal $\boldsymbol{r}$ is periodic with period K_r, i.e., $\exists\, k'$ such that $\boldsymbol{r}[k + K_\mathrm{r}] = \boldsymbol{r}[k]$, $\forall k \geqslant k'$.* ◊

Whilst the former might seem restrictive (see also our discussion), note that periodic references are common in applications such as robot manipulators, mechatronic rotary systems, and power plants [22]; see also *repetitive control*.

We consider a *weak Byzantine adversary*: an adversary that has no system knowledge but has both disclosure resources and disruption capabilities of the measurements [6]. Furthermore, we define the information available to the adversary at time k, which is the combination of system knowledge and disclosure resources, as the set $\mathbb{I}_\mathrm{a}[k]$.

Assumption 3. *The information $\mathbb{I}_\mathrm{a}[k]$ available to the adversary $\mathcal{A}$ satisfies $\mathbb{I}_\mathrm{a}[k] \supseteq \{k_0, \ldots, k_i \wedge \boldsymbol{y}_0, \ldots, \boldsymbol{y}_i\}$ at time $k_i \leqslant k < k_{i+1}$, i.e. it contains at least the first $i+1$ transmitted measurements $\boldsymbol{y}_i$ and sample times k_i.* ◊

The parameter Δi in (8) needs to be correctly chosen in order to avoid detection. Specifically, given a periodic reference to be tracked with period K_r, the loop length must be approximately equal to an integer multiple of the period of the reference signal. Note that an adversary can (approximately) recover K_r from the observer outputs $\boldsymbol{y}_0, \ldots, \boldsymbol{y}_i$, and κ_i can be recovered from the past sample times k_i. With the information $\mathbb{I}_\mathrm{a}[k]$, and given an arbitrary number of cycles $N_\mathrm{a} \in \mathbb{N}$, the adversary chooses

$$\Delta_i = \arg\min_{i \in \mathbb{N}} |N_\mathrm{a} \cdot K_\mathrm{r} - (k_{I_\mathrm{a}} - k_{I_\mathrm{a}-i})|. \tag{9}$$

Evidently, (9) states that the adversary chooses Δ_i such that that the sum of the inter-sample times is closest to the desired integer multiple of the reference period.

In order to detect anomalies such as replay attacks, a detector $\mathcal{D}$ is introduced, see §III-B, which generates a scalar detection signal g_i. A binary hypothesis test, given by

$$g_i \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta, \qquad \begin{aligned} &\mathcal{H}_0 : \text{Nominal system operation}, && (10\text{a}) \\ &\mathcal{H}_1 : \text{System under attack}, && (10\text{b}) \end{aligned}$$

is used to determine the presence of an anomaly. Here, $\eta > 0$ is a detection threshold to be designed, and $h_i \in \{\mathcal{H}_0, \mathcal{H}_1\}$ denotes whether or not an alarm is raised at the i-th sample instant. Finally, let $p_\mathrm{fp}, p_\mathrm{fn} \in (0,1)$ denote the false (positive) alarm rate (which is a design parameter), and the missed detection rate (i.e. false negative rate), respectively:

$$p_\mathrm{fp} = \mathbb{P}[g_i > \eta \,|\, \mathcal{H}_0], \qquad p_\mathrm{fn} = \mathbb{P}[g_i < \eta \,|\, \mathcal{H}_1]. \tag{11}$$

The objective and constraints of the adversary are to be *disruptive*, i.e., force at some finite $k > 0$ that $\boldsymbol{x}[k] \notin \mathbb{X}_\mathrm{s}$, while remaining *stealthy*, i.e., $g_i < \eta$ for the entire duration of the attack. We can formally state the problem we address:

Problem statement. *Design a detector $\mathcal{D}$ and STC policy $\mathcal{S}$ capable of detecting stealthy replay attacks (which remain undetected by passive detection methods).*

III. Watermarking and Detection

In this section, we discuss the design of both the detector $\mathcal{D}$, as well as our novel STC watermarking policy $\mathcal{S}$.

A. χ^2 detector

As a detection mechanism, we propose the following (static) χ^2 detector $\mathcal{D}$:

$$\mathcal{D}: \quad \begin{aligned} \boldsymbol{z}_i &= \boldsymbol{y}_i^{\downarrow} - \boldsymbol{C}\hat{\boldsymbol{x}}_{i\,|\,i-1}, && (12\text{a}) \\ g_i &= \boldsymbol{z}_i^\mathrm{T} (\boldsymbol{C}\boldsymbol{\Sigma}_{\mathrm{x},i\,|\,i-1}\boldsymbol{C}^\mathrm{T} + \boldsymbol{\Sigma}_\mathrm{v})^{-1} \boldsymbol{z}_i, && (12\text{b}) \end{aligned}$$

where $\boldsymbol{z}_i$ is the residual and g_i is the detection signal which is fed through a binary hypothesis test as in (10). The above detection scheme is inspired by the one proposed in [15] and here extended to the case of aperiodic sampling.

One can verify that the detection signal g_i in (12b) follows a χ^2 distribution with n_y degrees of freedom under nominal system operation, which stems from the fact that the detector residual $\boldsymbol{z}_i \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma}_{\mathrm{z},i})$ [12], with $\boldsymbol{\Sigma}_{\mathrm{z},i}$ as in (7c). Calculating the threshold η based on a false alarm rate $p_\mathrm{fp} \in (0,1)$ is straightforward, and given by [23, Lemma 1]

$$\eta = 2 \cdot P^{-1}\Big(\frac{n_y}{2}, 1 - p_\mathrm{fp}\Big). \tag{13}$$

Here, P^{-1} denotes the inverse regularized lower incomplete gamma function. Whilst, in general, it is desirable to choose the false alarm rate p_fp small, one cannot make it arbitrarily small without also increasing the missed detection rate.

B. Self-triggering policy

Let us denote by $\vec{\boldsymbol{x}}_i[\kappa] := \boldsymbol{A}_\kappa \hat{\boldsymbol{x}}_i + \boldsymbol{B}_\kappa \boldsymbol{u}_i$ the open-loop future state estimate, with matrices $\boldsymbol{A}_\kappa$, $\boldsymbol{B}_\kappa$ as in (5), and define $\boldsymbol{q}_i[\kappa] := \mathrm{col}(\tilde{\boldsymbol{x}}_i[\kappa], \hat{\boldsymbol{x}}_i, \boldsymbol{r}[k_i + \kappa])$. Let $\boldsymbol{Q} \in \mathbb{R}^{2 \cdot n_\mathrm{x} + n_\mathrm{y}}$ be a (quadratic) *triggering matrix* to be designed and define the function $\Gamma : \mathbb{R}^{n_\mathrm{x}} \times \mathbb{R}^{n_\mathrm{y}} \times \mathbb{R}^{n_\mathrm{u}} \to \mathbb{N}$ as

$$\Gamma(\hat{\boldsymbol{x}}_i, \boldsymbol{r}, \boldsymbol{u}_i) = \max_\kappa \{\kappa \leqslant \bar{\kappa} \,|\, \boldsymbol{q}_i^\mathrm{T}[\kappa] \boldsymbol{Q} \boldsymbol{q}_i[\kappa] \leqslant 0\}, \tag{14}$$

determining the next sampling instant. As is common in STC literature [2], we introduce an upper bound $\bar{\kappa} \in \mathbb{N}$ on the time between sampling instants. We assume the following:

Assumption 4. *The triggering matrix $\boldsymbol{Q}$ is designed such that any STC policy $\mathcal{S}$ producing triggering times*

$$\kappa_{i+1} \leqslant \Gamma(\hat{\boldsymbol{x}}_i, \boldsymbol{r}, \boldsymbol{u}_i), \qquad \forall \hat{\boldsymbol{x}}_i, \boldsymbol{r}, \boldsymbol{u}_i, \tag{15}$$

guarantees practical mean square stable (MSS) tracking of the reference signal $\boldsymbol{r}$ by system (1)-(3). That is, there exists a class-$\mathcal{KL}$ function[2] *β and constant $\gamma > 0$, such that*

$$\mathbb{E}\big[\|\boldsymbol{y}[k] - \boldsymbol{r}[k]\|^2\big] \leqslant \beta\big(\mathbb{E}\big[\|\boldsymbol{y}[0] - \boldsymbol{r}[0]\|^2\big], k\big) + \gamma, \tag{16}$$

for all $\boldsymbol{y}[0]$, $\boldsymbol{r} \in \mathbb{R}^\mathrm{y}$. ◊

Remark III.1. *Early triggering has been used before [24], [2], coming from the intuition that for quadratic Lyapunov-based triggering conditions, early triggering preserves stability. For bounded disturbances, sufficient conditions in terms of $\mathcal{L}_2$-stability were derived in [25]. In [26] an STC triggering policy that guarantees MSS is provided, which is preserved with early triggering.*

C. Early triggering watermarking

From here on, we will use the shorthand notation $\bar{\kappa}_{i+1} := \Gamma(\hat{\boldsymbol{x}}_i, \boldsymbol{r}, \boldsymbol{u}_i)$ to denote the next sampling *deadline* [24]. We employ a non-deterministic STC policy $\mathcal{S}$, which acts as the watermarking in our proposed method:

$$\mathcal{S}: \quad k_{i+1} = k_i + \kappa_{i+1}, \quad \kappa_{i+1} \sim p_{i+1}(\hat{\boldsymbol{x}}_i, \boldsymbol{u}_i), \tag{17}$$

[2] A continuous function $\beta : [0, a) \times [0, \infty) \to \mathbb{R}_{\geqslant 0}$ belong to class $\mathcal{KL}$ if, for all fixed s, $\beta(0, s) = 0$ and $\beta(r, s) > \beta(r', s)$, $\forall r > r'$, and, for all fixed r, $\beta(r, s) < \beta(r, s')$, $\forall s > s'$, and if $\lim_{s\to\infty} \beta(r, s) = 0$.

where p_{i+1} is a discrete probability distribution of the next inter-sample time, with $\kappa_{i+1} \in \{1, \ldots, \bar{\kappa}_{i+1}\}$. The entries $p_{i,\kappa} := p_{i+1}(\kappa \,|\, \hat{\boldsymbol{x}}_i, \boldsymbol{u}_i)$ need, in general, to be designed online at each sample time k_i. The open-loop future output estimate $\vec{\boldsymbol{y}}_i[\kappa] := \boldsymbol{C}\vec{\boldsymbol{x}}_i[\kappa]$ depends on κ, and follows a normal distribution $\mathcal{N}_\kappa$ with mean and covariance matrix given by

$$\boldsymbol{\mu}_\kappa(\hat{\boldsymbol{x}}_i, \boldsymbol{u}_i) = \boldsymbol{C}(\boldsymbol{A}_\kappa \hat{\boldsymbol{x}}_i + \boldsymbol{B}_\kappa \boldsymbol{u}_i), \tag{18a}$$

$$\boldsymbol{\Sigma}_\kappa(\boldsymbol{\Sigma}_{\mathrm{x},i}) = \boldsymbol{C}(\boldsymbol{A}_\kappa \boldsymbol{\Sigma}_{\mathrm{x},i} \boldsymbol{A}_\kappa^{\mathrm{T}} + \boldsymbol{\Sigma}_{\mathrm{w},\kappa})\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{\Sigma}_{\mathrm{v}}. \tag{18b}$$

Consider two event times k_i and $k_{i'}$, $k_i > k_{i'}$, with $\hat{\boldsymbol{x}}_i = \hat{\boldsymbol{x}}_{i'}$, $\boldsymbol{r}[k_i] = \boldsymbol{r}[k_{i'}]$, and $\boldsymbol{u}_i = \boldsymbol{u}_{i'}$, so that $\Gamma(\hat{\boldsymbol{x}}_i, \boldsymbol{r}, \boldsymbol{u}_i) = \Gamma(\hat{\boldsymbol{x}}_{i'}, \boldsymbol{r}, \boldsymbol{u}_{i'})$. An attacker in this situation may replay, at $k_i + \kappa$, the measurement $\boldsymbol{y}^{\downarrow}_{i+1} = \boldsymbol{y}_{i'+1}$ recorded at $k_{i'} + \kappa'$. However, if the next inter-sample time κ, prescribed by the STC mechanism $\mathcal{S}$, is not equal to κ', the open-loop future residual estimate $\vec{\boldsymbol{z}}_{i+1} = \boldsymbol{y}^{\downarrow}_{i+1} - \vec{\boldsymbol{y}}_i[\kappa]$ satisfies

$$\mathbb{E}\big[\vec{\boldsymbol{z}}_{i+1}\big] = \boldsymbol{\mu}_{\kappa'}(\hat{\boldsymbol{x}}_{i'}, \boldsymbol{u}_{i'}) - \boldsymbol{\mu}_\kappa(\hat{\boldsymbol{x}}_i, \boldsymbol{u}_i) =: \boldsymbol{\mu}_{\mathrm{z},i}[\kappa, \kappa'], \tag{19}$$

which, even in this most beneficial case for the adversary, that is identical initial conditions ($\hat{\boldsymbol{x}}_i = \hat{\boldsymbol{x}}_{i'}$, $\boldsymbol{r}[k_i] = \boldsymbol{r}[k_i']$, $\boldsymbol{u}_i = \boldsymbol{u}_{i'}$), is non-zero if $\kappa' \neq \kappa$. We leverage this observation to design an STC policy $\mathcal{S}$ capable of detecting the attack. In particular, we exploit the fact that the distribution of $\boldsymbol{z}_i$ (and therefore g_i) during a replay attack can be tuned through the design of p_{i+1}, as we have

$$\mathcal{H}_0: \qquad \mathbb{E}[\boldsymbol{z}_i] = \boldsymbol{0}, \tag{20a}$$

$$\mathcal{H}_1: \qquad \mathbb{E}[\boldsymbol{z}_i] = \sum_{\kappa,\kappa'=1}^{\bar{\kappa}_{i+1}} p_{i,\kappa} \cdot p_{i,\kappa'} \cdot \boldsymbol{\mu}_{\mathrm{z},i-1}[\kappa, \kappa']. \tag{20b}$$

Ideally, we would like to design the distribution p_{i+1} such that the missed detection rate p_{fn} is minimized. The former is achieved when, during a replay attack, there is a high probability of a mismatch between κ (the next inter-event time) and κ' (the one recorded by the attacker). However, computing the missed detection rate p_{fp} in closed-form is hard [15]. We propose to design p_i by solving at each event instant i:

$$\min_{\boldsymbol{p}_i} \boldsymbol{p}_i^{\mathrm{T}} \boldsymbol{W}_{\mathrm{p}}(\hat{\boldsymbol{x}}_i, \boldsymbol{u}_i) \boldsymbol{p}_i \quad \text{s.t.} \quad \|\boldsymbol{p}_i\|_1 = 1, \quad \boldsymbol{p}_i \geqslant 0, \tag{21}$$

with $\boldsymbol{p}_i := \mathrm{col}(p_{i,1}, \ldots, p_{i,\bar{\kappa}_{i+1}}) \in \mathbb{R}^{\bar{\kappa}_{i+1}}$ as the decision variable, and $\boldsymbol{W}_{\mathrm{p}}(\hat{\boldsymbol{x}}_i, \boldsymbol{u}_i) \succ 0$ a matrix to be designed.

Inspired by [15], we propose an approach based on incorporating the Kullback-Leibler (KL) divergence between the distribution of the open-loop future output estimates (18), for different inter-event times κ, κ', into the matrix $\boldsymbol{W}_{\mathrm{p}}$. The early triggering aims to minimize the missed detection rate during a replay attack. The use of the KL divergence is motivated by the fact that minimizing the missed detection is related to maximizing the KL divergence between the distributions of $\vec{\boldsymbol{y}}_i[\kappa]$ and $\vec{\boldsymbol{y}}_i[\kappa']$ for $\kappa \neq \kappa'$. For two normal distributions $\mathcal{N}_\kappa$ and $\mathcal{N}_{\kappa'}$, the KL divergence is known in closed form and can be readily computed [27]. As the KL divergence is in general not symmetric, we opt for the use of the (symmetric) Jeffrey's divergence instead, given by $D_{\mathrm{J}}(\mathcal{N}_\kappa \| \mathcal{N}_{\kappa'}) = D_{\mathrm{KL}}(\mathcal{N}_\kappa \| \mathcal{N}_{\kappa'}) + D_{\mathrm{KL}}(\mathcal{N}_{\kappa'} \| \mathcal{N}_\kappa)$. Finally, to prevent $\boldsymbol{W}_{\mathrm{p}}$ from not being (strictly) positive definite, we define $\boldsymbol{W}_{\mathrm{p}} = \tilde{\boldsymbol{W}}_{\mathrm{p}} + |\min\{0, \lambda_{\min}(\tilde{\boldsymbol{W}}_p)\} - \epsilon_0| \cdot \boldsymbol{I}$ for some small positive $\epsilon_0 \approx 0$, where $\tilde{\boldsymbol{W}}_{\mathrm{p}} \in \mathbb{R}^{\bar{\kappa}_{i+1} \times \bar{\kappa}_{i+1}}$ has entries

$$\tilde{w}_{\mathrm{p},ij} = e^{-\sqrt{p_{\mathrm{fp}}} \cdot D_{\mathrm{J}}(\mathcal{N}_i \| \mathcal{N}_j)}. \tag{22}$$

Whilst (21) needs to be solved online at each sample time k_i, this does not appear to be prohibitive, considering it is a QP with a relatively low number of decision variables and constraints.

Remark III.2. *Additionally, under mild assumptions, whenever $\boldsymbol{u}$ is "large" compared to $\boldsymbol{w}$, $\boldsymbol{v}$, the optimal solution to (21) becomes a discrete uniform distribution [28, Proposition 5.2]). Thus, if the above holds for all inputs (which can be checked* a postiori*), the proposed early triggering mechanism can be fully designed offline.*

The proposed design is exemplified in Fig. 2.

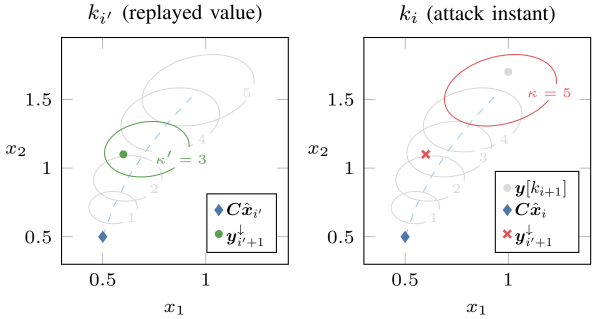

Fig. 2. Visualization of the proposed method, with ellipses given by $\{\boldsymbol{y} \in \mathbb{R}^{n_{\mathrm{y}}} \,|\, (\boldsymbol{y} - \vec{\boldsymbol{y}}_i[\kappa])^{\mathrm{T}} \boldsymbol{\Sigma}_\kappa(\boldsymbol{\Sigma}_{\mathrm{x},i})(\boldsymbol{y} - \vec{\boldsymbol{y}}_i[\kappa]) = \eta\}$: their exteriors depict when an alarm is triggered for that particular κ. The dashed line denotes $\vec{\boldsymbol{y}}_i[\kappa]$, both for $\kappa \in \{1 \ldots, 5\}$. If, during a replay attack, at time k_{i+1}, the measurement $\boldsymbol{y}_{i'+1}$ is replayed but $\kappa' = 3 \neq \kappa = 5$, an alarm is raised and the attack is detected.

Remark III.3. *Our design methodology is similar to an* emulation-based *approach [1], where first a controller $\mathcal{C}$ is designed such that stability guarantees and desired performance criteria are met using periodic sampling. Then, an STC policy is designed where performance is traded off for fewer communications. Finally, the STC policy is augmented with early triggering to allow for attack detection. As such, our proposed design methodology is applicable to legacy systems (where the control logic has already been designed).*

IV. Illustrative Example

Consider the discrete-time unstable plant $\mathcal{P}$ given by:

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0.1 \\ 0.035 & 0.99 \end{bmatrix}, \qquad \boldsymbol{B} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \tag{23}$$

and $\boldsymbol{C} = \boldsymbol{I}$ [29]. Furthermore, consider the digital controller $\mathcal{C}$ given by $\boldsymbol{A}_{\mathrm{c}} = \boldsymbol{I}$, $\boldsymbol{B}_{\mathrm{c}} = 0.1 \cdot \boldsymbol{I}$, and the matrices

$$\boldsymbol{C}_{\mathrm{c}} = \begin{bmatrix} 0.01 & 0.022 \end{bmatrix}, \quad \boldsymbol{D}_{\mathrm{c}} = \begin{bmatrix} 0.0875 & 0.198 \end{bmatrix}. \tag{24}$$

The objective of the closed-loop control system is to track a sinusoidal reference signal given by $\boldsymbol{r}[k] = \mathrm{col}(A_{\mathrm{r}} \cdot$

$\sin(2\pi \cdot k/K_\mathrm{r}), A_\mathrm{r} \cdot \cos(2\pi \cdot k/K_\mathrm{r}))$, with amplitude $A_\mathrm{r} = 2$ and period $K_\mathrm{r} = 60$. Since Assumption 2 holds, the control system is susceptible to a replay attack. The process and measurement noise covariance matrices are set to $\boldsymbol{\Sigma}_\mathrm{w} = 10^{-3} \cdot \boldsymbol{I}$ and $\boldsymbol{\Sigma}_\mathrm{v} = 10^{-4} \cdot \boldsymbol{I}$, respectively.

We consider four scenarios: the first is with a 'greedy' STC policy $\bar{\mathcal{S}}$ such that $\kappa_{i+1} = \bar{\kappa}_{i+1}$, without any additional countermeasures. We refer to this scenario as the *baseline*. Next, we consider our proposed method with STC policy $\mathcal{S}$ from (17), as discussed in §III-B, and finally, we consider two scenarios with STC policy $\bar{\mathcal{S}}$ and additive [15] and multiplicative [29] watermarking, respectively.

The STC policies $\bar{\mathcal{S}}, \mathcal{S}$ are designed with $\boldsymbol{Q}$ as

$$\boldsymbol{Q} = \begin{bmatrix} \boldsymbol{I} - \sigma^2 \cdot \boldsymbol{C}^\mathrm{T}\boldsymbol{C} & -\boldsymbol{I} & \sigma^2 \cdot \boldsymbol{C} \\ -\boldsymbol{I} & \boldsymbol{I} & \boldsymbol{0} \\ \sigma^2 \cdot \boldsymbol{C}^\mathrm{T} & \boldsymbol{0} & -\sigma^2 \cdot \boldsymbol{I} \end{bmatrix}, \qquad (25)$$

and $\sigma = 0.32$, $\bar{\kappa} = 10$. From our simulation results, we find that Assumption 4 holds for all four scenarios.

Finally, the event-triggered χ^2 detector is designed with a false alarm rate $p_\mathrm{fp} = 0.1\%$. As $n_\mathrm{y} = 2$ this implies $\eta = 13.82$ from (13). The safe region was chosen as $\mathbb{X}_\mathrm{s} = \{\boldsymbol{x}_0 \in \mathbb{R}^{n_\mathrm{x}} \mid \|\boldsymbol{x}_0\|_\infty \leqslant 4\}$ such that, prior to any attack, $\boldsymbol{x}[k] \in \mathbb{X}_\mathrm{s}$ for all four considered scenarios.

We perform a simulation for 360 time steps, where at $k = 240$, the adversary $\mathcal{A}$ launches a replay attack with a delay of two cycles, meaning $N_\mathrm{a} = 2$. Given Assumption 3 and using (9), the adversary chooses $\Delta i = 41$ for the baseline and the two benchmarks, and $\Delta i = 52$ for our proposed method (as early triggering leads to more events per tracking period K_r).

In Fig. 3-5, the orange and red vertical lines denote k_{I_a} and the smallest k for which $\boldsymbol{x}[k] \notin \mathbb{X}_\mathrm{s}$, respectively. The dashed line in the top plot depicts the genuine measurement $\boldsymbol{y}[k]$, unavailable to the controller.

In Fig. 3, the results for the baseline scenario are depicted. Observe that even when $\boldsymbol{x}[k]$ leaves $\mathbb{X}_\mathrm{s}$, no alarm is raised, demonstrating that the passive detection scheme is inadequate. On the other hand, our proposed approach, depicted in Fig. 4, detects the replay attack before the state trajectory leaves the safe region.

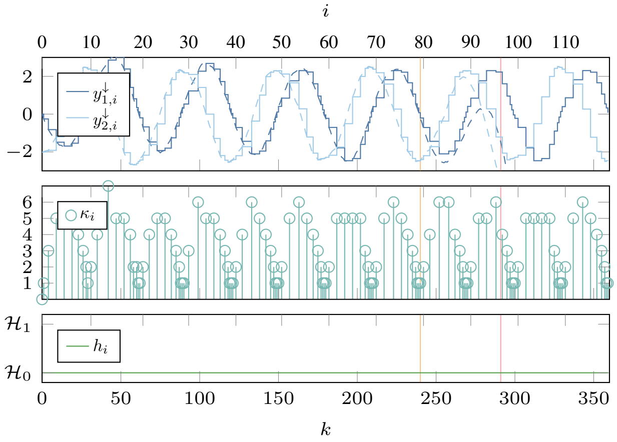

Fig. 3. Results of the baseline.

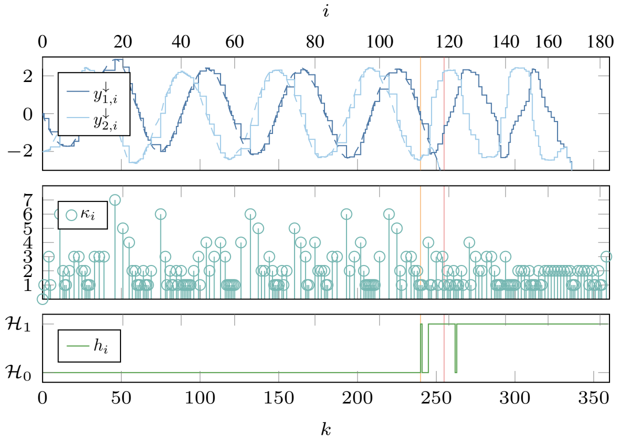

Fig. 4. Performance of our proposed method.

Next, we consider the two watermarking benchmarks. To provide a fair comparison with additive watermarking, the input noise covariance matrix $\boldsymbol{\Sigma}_\mathrm{u}$ (see [10, 23]) is set to $\boldsymbol{\Sigma}_\mathrm{u} = \delta \cdot \boldsymbol{I}$, with $\delta \in 10^{-3} \cdot \mathbb{N}$ as the smallest value such that the replay attack is detected before the state trajectory leaves the safe region $\mathbb{X}_\mathrm{s}$ (note that the procedure as in [15] is not applicable, as we are considering a dynamic controller). For the scenario considered, we find $\delta = 2.7 \cdot 10^{-2}$. Lastly, for multiplicative watermarking, we choose a filter order $N_\mathrm{w} = 4$ and switching instants $k \in 120 \cdot \mathbb{N}_0$, similar to [29]. The results can be seen in Fig. 5, where we observe that both benchmarks detect the replay attack before $\boldsymbol{x}[k] \notin \mathbb{X}_\mathrm{s}$.

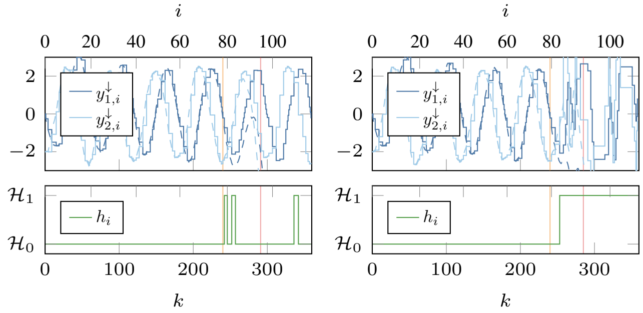

Fig. 5. Left: additive watermarking, right: multiplicative watermarking.

To compare the nominal performance of our proposed method as well as the two benchmarks to the baseline, we let the simulation run in the attack-free case for $K = 10^5$ time steps (resulting in a different number of events N_i for each scenario). As metrics, we use the average inter-event time $\kappa_\mathrm{avg} = 1/N_i \cdot \sum_{i=0}^{N_\mathrm{i}} \kappa_i$, the average (maximum) squared tracking error $E_\mathrm{avg} = 1/K \cdot \sum_{k=0}^{K} \|\boldsymbol{r}[k] - \boldsymbol{y}[k]\|^2$ ($E_\mathrm{max} = \max_k \|\boldsymbol{r}[k] - \boldsymbol{y}[k]\|^2$), and the average squared change in actuation $\Delta U_\mathrm{avg} = 1/(K-1) \cdot \sum_{k=1}^{K} \|\boldsymbol{u}[k] - \boldsymbol{u}[k-1]\|^2$. The result can be seen in Tab. I, where multiplicative watermarking is omitted as all metrics are identical to the baseline. From Tab. I, it can be seen that there exists a trade-off, as our proposed watermarking scheme requires more communications on average, but on the other hand, decreased the average actuation 'jitter' (caused by additive watermarking).

TABLE I

TRADE-OFFS BETWEEN THE PROPOSED METHOD AND BENCHMARK

	κ_{avg}	E_{avg}	$E_{\max}$	$\Delta U_{\mathrm{avg}} \cdot 10^3$
I	3.09	9.41	4.86	2.5
II	2.16 **−30.0%**	9.16 −3.1%	4.87 −4.7%	2.0 **−18.4%**
III	3.12 +0.8%	9.45 +0.4%	5.03 +4.9%	3.5 **+41.6%**

I: Baseline, II: Proposed method, III: Additive watermarking
Red: adverse increase/decrease. **Green**: beneficial increase/decrease.

Remark IV.1. *A key difference between our proposed method and multiplicative watermarking is that our scheme is entirely located on the controller side, whilst multiplicative adds additional computational load on the sensors. Additionally, the latter is susceptible to desynchronization [17], and (due to the watermarking filter) leads to a loss of* availability.

Furthermore, for additive watermarking, injecting noise into the control signal can burden the physical actuators [29] and lead to increased actuator attrition. Whilst our proposed method does lead to more transmissions compared to the baseline, no control performance is sacrificed: in fact, recomputing $\boldsymbol{u}_i$ more frequently leads to performance gains.

V. CONCLUSIONS AND FUTURE WORK

We presented a novel watermarking scheme capable of detecting replay attacks based on modifying an existing STC policy, making the scheme modular and applicable to legacy systems. A comparison with additive and multiplicative watermarking in an illustrative example showed that there exist trade-offs between competing performance criteria.

One of our main priorities is to provide design methodologies (e.g. extending the results in [26]) guaranteeing that Assumption 4 holds *a priori*. We are currently working on improving the design methodology such that the detection performance on regulation tasks remains adequate (see [28, §5.6]). Finally, the effects of network-induced phenomena such as communication delays and package drops on the detection scheme are also the subject of future research.

REFERENCES

[1] W. P. M. H. Heemels, M. C. F. Donkers, and A. R. Teel, "Periodic Event-Triggered Control for Linear Systems," *IEEE Trans. on Automatic Control*, vol. 58, no. 4, pp. 847–861, Apr. 2013.

[2] M. Mazo Jr, A. Anta, and P. Tabuada, "An ISS self-triggered implementation of linear controllers," *Automatica*, vol. 46, no. 8, pp. 1310–1314, Aug. 2010.

[3] R. M. Lee, M. J. Assante, and T. Conway, "Analysis of the cyber attack on the Ukrainian power grid," *SANS Industrial Control Systems*, 2016.

[4] H. Sandberg, "Cyber-Physical Security," in *Encyclopedia of Systems and Control*, J. Baillieul and T. Samad, Eds. London: Springer, 2020, pp. 1–8.

[5] A. Teixeira, I. Shames, H. Sandberg, and K. H. Johansson, "Revealing stealthy attacks in control systems," in *2012 50th Annual Allerton Conf. on Communication, Control, and Computing (Allerton)*, Oct. 2012, pp. 1806–1813.

[6] ——, "A secure control framework for resource-limited adversaries," *Automatica*, vol. 51, pp. 135–148, Jan. 2015.

[7] H. Sandberg, V. Gupta, and K. H. Johansson, "Secure Networked Control Systems," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 5, no. 1, pp. 445–464, 2022.

[8] A. Teixeira, S. Amin, H. Sandberg, K. H. Johansson, and S. S. Sastry, "Cyber security analysis of state estimators in electric power systems," in *49th IEEE Conf. on Decision and Control (CDC)*. IEEE, 2010, pp. 5991–5998.

[9] C. Kwon, W. Liu, and I. Hwang, "Security analysis for cyber-physical systems against stealthy deception attacks," in *2013 American Control Conf.* IEEE, 2013, pp. 3344–3349.

[10] Y. Mo and B. Sinopoli, "Secure control against replay attacks," in *2009 47th Annual Allerton Conf. on Communication, Control, and Computing (Allerton)*, Sep. 2009, pp. 911–918.

[11] H. Liu, J. Yan, Y. Mo, and K. H. Johansson, "An On-line Design of Physical Watermarks," Sep. 2018.

[12] B. Yaghooti, R. Romagnoli, and B. Sinopoli, "Physical watermarking for replay attack detection in continuous-time systems," *European Journal of Control*, vol. 62, pp. 57–62, Nov. 2021.

[13] R. Romagnoli, S. Weerakkody, and B. Sinopoli, "A Model Inversion Based Watermark for Replay Attack Detection with Output Tracking," in *2019 American Control Conf. (ACC)*, Jul. 2019, pp. 384–390.

[14] F. Miao, Q. Zhu, M. Pajic, and G. J. Pappas, "Coding schemes for securing cyber-physical systems against stealthy data injection attacks," *IEEE Trans. on Control of Network Systems*, vol. 4, no. 1, pp. 106–117, 2016.

[15] Y. Mo, S. Weerakkody, and B. Sinopoli, "Physical Authentication of Control Systems: Designing Watermarked Control Inputs to Detect Counterfeit Sensor Outputs," *IEEE Control Systems Magazine*, vol. 35, no. 1, pp. 93–109, Feb. 2015.

[16] P. Griffioen, S. Weerakkody, and B. Sinopoli, "A moving target defense for securing cyber-physical systems," *IEEE Trans. on Automatic Control*, vol. 66, no. 5, pp. 2016–2031, 2020.

[17] R. M. G. Ferrari and A. M. H. Teixeira, "A Switching Multiplicative Watermarking Scheme for Detection of Stealthy Cyber-Attacks," *IEEE Trans. on Automatic Control*, vol. 66, no. 6, pp. 2558–2573, Jun. 2021.

[18] V. S. Dolk, P. Tesi, C. De Persis, and W. P. M. H. Heemels, "Event-Triggered Control Systems Under Denial-of-Service Attacks," *IEEE Trans. on Control of Network Systems*, vol. 4, no. 1, pp. 93–105, Mar. 2017.

[19] C. De Persis and P. Tesi, "Resilient Control under Denial-of-Service," *IFAC Proceedings Volumes*, vol. 47, no. 3, pp. 134–139, Jan. 2014.

[20] D. Du, C. Zhang, X. Li, M. Fei, and H. Zhou, "Attack Detection for Networked Control Systems Using Event-Triggered Dynamic Watermarking," *IEEE Trans. on Industrial Informatics*, vol. 19, no. 1, pp. 351–361, Jan. 2023.

[21] A. Barboni, A. W. Al-Dabbagh, and T. Parisini, "An Event-Triggered Watermarking Strategy for Detection of Replay Attacks," *IFAC-PapersOnLine*, vol. 55, no. 6, pp. 317–322, Jan. 2022.

[22] G. Ma, X. Liu, P. R. Pagilla, and X. Yu, "Two-Channel Periodic Event-Triggered Observer-Based Repetitive Control for Periodic Reference Tracking," in *IECON 2018 - 44th Annual Conf. of the IEEE Industrial Electronics Society*. Washington, District of Columbia, USA: IEEE, Oct. 2018, pp. 2469–2474.

[23] N. Hashemi and J. Ruths, "Generalized chi-squared detector for LTI systems with non-Gaussian noise," in *IEEE Int. Conf. on Cyber Technology in Automation*. Philadelphia: IEEE, Jul. 2019, pp. 404–410.

[24] G. d. A. Gleizer, K. Madnani, and M. Mazo Jr, "Self-Triggered Control for Near-Maximal Average Inter-Sample Time," in *60th IEEE Conf. on Decision and Control (CDC)*, Dec. 2021, pp. 1308–1313.

[25] G. d. A. Gleizer and M. Mazo Jr, "Self-Triggered Output Feedback Control for Perturbed Linear Systems," *IFAC-PapersOnLine*, vol. 51, no. 23, pp. 248–253, Jan. 2018.

[26] R. P. Anderson, D. Milutinović, and D. V. Dimarogonas, "Self-triggered sampling for second-moment stability of state-feedback controlled SDE systems," *Automatica*, vol. 54, pp. 8–15, Apr. 2015.

[27] D. Belov and R. Armstrong, "Distributions of the Kullback-Leibler divergence with applications," *The British journal of mathematical and statistical psychology*, vol. 64, pp. 291–309, May 2011.

[28] B. Wolleswinkel, "A secure control framework for self-triggered control: Exploiting aperiodic sampling for the detection and prevention of stealthy attacks," Master's thesis, Delft University of Technology, Delft, Feb. 2024.

[29] R. M. G. Ferrari and A. M. H. Teixeira, "Detection and Isolation of Replay Attacks through Sensor Watermarking," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 7363–7368, Jul. 2017.