Authentication of Cyber-Physical Systems Under Learning-based Attacks

Mohammad Javad Khojasteh, Anatoly Khina, Massimo Franceschetti, and Tara Javidi

Abstract—The problem of attacking and authenticating cyber-physical systems is considered. This paper concentrates on the case of a scalar, discrete-time, time-invariant, linear plant under an attack which can override the sensor and the controller signals. Prior works assumed the system was known to all parties and developed watermark-based methods. In contrast, in this paper the attacker needs to learn the open-loop gain in order to carry out a successful attack. A class of two-phase attacks is considered: during an exploration phase, the attacker passively eavesdrops and learns the plant dynamics, followed by an exploitation phase, during which the attacker hijacks the input to the plant and replaces the input to the controller with a carefully crafted fictitious sensor reading with the objective of detecting a possible attack. The attacker, an authentication test that examines the variance over a time window, tools from information theory and statistics are utilized to derive bounds on the detection and deception probabilities with and without a watermark signal, when the attacker uses an arbitrary learning algorithm to estimate the open-loop gain of the plant.

Index Terms—Cyber-physical systems security, secure control, physical authentication of control systems, man-in-the-middle attack, system identification

I. INTRODUCTION

The recent technological advances in wireless communications and computation, and their integration into networked control and cyber-physical systems (CPS) [1], open the door to a myriad of new and exciting opportunities in transportation, health care, agriculture, energy, and many others.

However, the distributed nature of CPS is often a source of vulnerability [2]–[4]. Security breaches in CPS can have catastrophic consequences ranging from hampering the economy by obtaining financial gain, through hijacking autonomous vehicles and drones, and all the way to terrorism by manipulating life-critical infrastructures [5]–[9]. Real-world instances of security breaches in CPS that were discovered and made available to the public include the revenge sewage attack in Maroochy Shire, Australia [10], the Ukraine power attack [11], the German steel mill cyber-attack [12] and the Iranian uranium enrichment facility attack via the Stuxnet malware [13]–[17]. Consequently, studying and preventing such security breaches via control-theoretic methods has received a great deal of attention in recent years [18]–[29].

An important and widely used class of attacks on CPS are based on the “man-in-the-middle” (MITM) attack technique (cf. [30]): an attacker takes over the physical plant’s control and sensor signals. The attacker overrides the control signals with malicious inputs in order to push the plant toward an alternative trajectory, often unstable and catastrophic. Consequently, many CPS constantly monitor/sense the plant outputs with the objective of detecting a possible attack. The attacker, on the other hand, aims to overwrite the sensor readings in a manner that would be indistinguishable from the legitimate ones.

The simplest instance a MITM attack is the replay attack [31]–[33], in which the attacker observes and records the legitimate system behavior across a long period of time and then replays it at the controller’s input; this attack is reminiscent of the notorious attack of video surveillance systems, in which previously recorded surveillance footage is replayed during a heist. A well-known example of this attack is that of the Stuxnet malware, which used an operating system vulnerability to enable a twenty-one seconds long replay attack during which the attacker has driven the speed of the centrifuges at a uranium enrichment facility toward excessively high and destructive speed levels [34]. The extreme simplicity of the replay attack, which can be implemented with zero knowledge of the system dynamics and sensors specification, has made it a popular and well-studied topic of research [31]–[33], [35]–[37].

In contrast to the replay attack, a paradigm that follows Shannon’s maxim of Kerckhoffs’s principle: “the enemy knows the system,” was considered by Satchidanandan and Kumar [38] and Ko et al. [39]. This assumes that the attacker has complete knowledge of the dynamics and parameters of the system, which allows the attacker to construct arbitrarily long fictitious sensor readings, that are statistically identical to the actual signals, without being detected.

To counter both the replay and the “statistical-duplicate” attacks, Mo and Sinopoli [32], and Satchidanandan and Kumar [38], respectively, proposed to superimpose a random watermark, unknown to the attacker, on top of the (optimal) control signal. In this way, by testing the correlation of the subsequent measurements with the watermark signal, the controller is able to detect the attack. By superimposing watermarking at different power levels, improved attack detection
probability can be traded for an increase in the control cost.

The two interesting approaches described above suffer from some shortcomings. First, in the case of a replay attack the usage of the watermarking signal is unnecessary: by taking a long enough detection window, the controller is always able to detect such an attack even in the absence of watermarks by simply testing for repetitions. A watermark is only necessary when the detection window of the controller is small compared to the recording (and replay) window of the attacker. Second, in the case of a statistical-duplicate attack, we must assume that the attacker has no access to the signal generated and applied by the controller. Since this type of attack assumes the attacker has full system knowledge, if it also has access to the control signal then it can construct a fictitious sensor readings containing any watermark signal inscribed by the controller. Assuming there is no access to the control signal seems a questionable assumption for an attacker who is capable of hijacking the whole system and overriding the control signal.

The two approaches constitute two extremes: the replay attack assumes no knowledge of the system parameters— and as a consequence it is relatively easy to detect. The statistical-duplicate attack assumes full knowledge of the system dynamics—and as a consequence it requires a more sophisticated detection procedure, as well as additional assumptions to ensure it can be detected.

In the current work, we explore a model that is somewhat in between these two extremes. We assume that only the controller has perfect knowledge of the system dynamics. This is a reasonable assumption, since the controller is tuned in between these two extremes. We assume that only the controller has perfect knowledge of the system dynamics to a far greater precision than the attacker. On the other hand, we assume the attacker knows that the system dynamics to a far greater precision than the attacker. Assuming there is no access to the control signal seems a questionable assumption for an attacker who is capable of hijacking the whole system and overriding the control signal.

Our approach is reminiscent of parametric linear system identification (SysID), but in contrast to classical SysID our attacker is constrained to passive identification. Specifically, we consider two-phase attacks akin to the exploration and exploitation phases in reinforcement learning/multi-armed bandit problems [41], [42]: in the exploration phase the attacker passively listens and learns the system parameter(s); in the exploitation phase the attacker uses the learned parameter(s) of the first phase to try and mimic the statistical behavior of the real plant, in a similar fashion to the statistical-duplicate attack. For the case of two-phase linear attacks, we analyze the achievable performance of a least-squares (LS) estimation-based scheme and a variance detection test. We also provide a lower bound on the attack-detection probability under the variance detection test and any learning algorithm. We provide explicit results for the case where the duration of the exploitation phase tends to infinity. To enhance the security of the system, we also extend the results to the case of a superimposed watermark (or authentication) signal.

An outline of the rest of the paper is as follows. We set up the problem in Sec.I and state the main results in Sec.III with their proofs relegated to the appendix. Simulations are provided in Sec.IV. We conclude with a discussion of future research directions in Sec.V.

A. Notation

We denote by \( \mathbb{N} \) the set of natural numbers. All logarithms, denoted by \( \log \), are to base 2. For two real valued functions \( g \) and \( h \), \( g(x) = O(h(x)) \) as \( x \to \infty \) means \( \lim_{x \to \infty} |g(x)/h(x)| < \infty \), and \( g(x) = o(h(x)) \) as \( x \to \infty \) means \( \lim_{x \to \infty} |g(x)/h(x)| = 0 \). We denote by \( X_i^j = (x_i, \ldots, x_j) \) the realization of the random vector \( X_i^j = (X_{i+1}, \ldots, X_j) \) for \( i, j \in \mathbb{N}, i \leq j \). \( \| \cdot \| \) denotes the Euclidean norm. \( \mathbb{P}_X \) denotes the distribution of the random variable \( X \) with respect to (w.r.t.) probability measure \( \mathbb{P} \), whereas \( f_X \) denotes its probability density function (PDF) w.r.t. to the Lebesgue measure, if it has one. An event happens almost surely (a.s.) if it occurs with probability one. For real numbers \( a \) and \( b \), \( a < b \) means \( a \) is much less than \( b \), in some numerical sense, while for probability distributions \( \mathbb{P} \) and \( \mathbb{Q} \), \( \mathbb{P} \ll \mathbb{Q} \) means \( \mathbb{P} \) is absolutely continuous w.r.t. \( \mathbb{Q} \). \( d\mathbb{P}/d\mathbb{Q} \) denotes the Radon–Nikodym derivative of \( \mathbb{P} \) w.r.t. \( \mathbb{Q} \). The Kullback–Leibler (KL) divergence between probability distributions \( \mathbb{P}_X \) and \( \mathbb{P}_Y \) is defined as

\[
D(\mathbb{P}_X \| \mathbb{P}_Y) \triangleq \begin{cases} 
\mathbb{E}_{\mathbb{P}_X} \left[ \log \frac{d\mathbb{P}_X}{d\mathbb{P}_Y} \right], & \mathbb{P}_X \ll \mathbb{P}_Y; \\
\infty, & \text{otherwise},
\end{cases}
\]

where \( \mathbb{E}_{\mathbb{P}_X} \) denotes the expectation w.r.t. probability measure \( \mathbb{P}_X \). The conditional KL divergence between probability distributions \( \mathbb{P}_{X|Y} \) and \( \mathbb{Q}_{Y|X} \) averaged over \( \mathbb{P}_X \) is defined as

\[
D \left( \mathbb{P}_{X|Y} \| \mathbb{Q}_{Y|X} \big| \mathbb{P}_X \right) \triangleq \mathbb{E}_{\mathbb{P}_X} \left[ D \left( \mathbb{P}_{X|Y} \| \mathbb{Q}_{Y|X} \right) \right],
\]

where \( (X, Y) \) are independent and identically distributed (i.i.d.). The mutual information between random variables \( X \) and \( Y \) is defined as \( I(X; Y) \triangleq D(\mathbb{P}_{X,Y} \| \mathbb{P}_X \mathbb{P}_Y) \). The conditional mutual information between random variables \( X \) and \( Y \) given random variable \( Z \) is defined as \( I(X; Y \| Z) \triangleq \mathbb{E}_{\mathbb{P}_X} \left[ I(X; Y \| Z = Z) \right] \), where \( (Z, \bar{Z}) \) are i.i.d.

II. PROBLEM SETUP

We consider the networked control system depicted in Fig.1, where the plant dynamics are described by a scalar, discrete-time, linear time-invariant (LTI) system

\[
X_{k+1} = aX_k + U_k + W_k,
\]

where \( X_k \), \( a \), \( U_k \), \( W_k \) are real numbers representing the plant state, open-loop gain of the plant, control input, and plant disturbance, respectively, at time \( k \in \mathbb{N} \). The controller, at time \( k \), observes \( Y_k \) and generates a control signal \( U_k \) as a function of \( Y_k \). We assume that the initial condition \( X_0 \) has a known (to all parties) distribution and is independent of the disturbance sequence \( \{W_k\} \), which is an i.i.d. process with PDF known to all parties. We assume that \( U_0 = W_0 = 0 \).
A is further assumed to be independent of \(X_0\) and \(\{W_k | k \in \mathbb{N}\}\).

We consider the (time-averaged) linear quadratic (LQ) control cost [43]:

\[
J_T(U) \triangleq \frac{1}{T} \mathbb{E}_{\mathcal{F}_n} \left[ \sum_{k=0}^{T} qX_k^2 + rU_k^2 \right],
\]

where the weights \(q\) and \(r\) are non-negative known (to the controller) real numbers that penalize the cost for state deviations and control actuations, respectively.

### A. Adaptive Integrity Attack

We define Adaptive Integrity Attacks (AIA) that consist of a passive and an active phases, referred to as exploration and exploitation, respectively.

During the exploration phase, depicted in Fig. 1a, the attacker eavesdrops and learns the system, without altering the input signal to the controller, i.e., \(Y_k = X_k\) during this phase.

On the other hand, during the exploitation phase, depicted in Fig. 1b, the attacker intervenes as a MITM in two different parts of the control loop with the aim of pushing the plant toward an alternative trajectory (usually unstable) without being detected by the controller: it hijacks the true measurements and feeds the controller with a fictitious input \(Y_k = V_k\) instead. Furthermore, it issues and overrides a malicious control signal to the actuator \(U_k\) instead of the signal \(U_k\) that is generated by the controller as depicted in Fig. 1b

**Remark 1.** Attacks that manipulate the control signal by tampering the integrity of the sensor readings, while trying to remain undetected, are usually referred to as integrity attacks, e.g., [32]. Since in the class of attacks described above, the attacker learns the open-loop gain of the plant in a fashion reminiscent of adaptive control techniques, we referred to attacks in this class as AIA.

### B. Two-Phase AIA

While in a general AIA the attacker can switch between the exploration and exploitation phases back and forth or try to combine them together in an online fashion (see Sec. V-A), in this work, we concentrate on a special class of AIA comprising only two disjoint consecutive phases as follows.

**Phase 1: Exploration.** As illustrated in Fig. 1a for \(k \in [0, L]\) the attacker observes the plant state and control input, and tries to learn the open-loop gain \(a\). We denote by \(\hat{A}\) the attacker’s estimate of the open-loop gain \(a\).

**Phase 2: Exploitation.** As illustrated in Fig. 1b from time \(L + 1\) onwards the attacker hijacks the system and feeds a malicious control signal to the plant \(\hat{U}_k\) and a fictitious sensor reading \(\hat{Y}_k = V_k\) to the controller.

### C. Linear Two-Phase AIA

A linear two-phase attack is a special case of the two-phase AIA of Sec. V-B in which the exploitation phase of the attacker takes the following linear form.

\[
V_{k+1} = \hat{A}V_k + U_k + \hat{W}_k, \quad k = L, \ldots, T - 1,
\]
where $\bar{W}_k$ for $k = L, \ldots, T - 1$ are i.i.d. with $f_{\bar{W}} = f_W$; $U_k$ is the control signal generated by the controller, which is fed with the fictitious virtual signal $V_k$ by the attacker; $V_L = X_L$; and $\hat{A}$ is the estimate of the open-loop gain of the plant at the conclusion of Phase 1.

The controller/detector, being aware of the dynamic (1) and the open-loop gain $a$, attempts to detect possible attacks by testing for statistical deviations from the typical behavior of the system (1). More precisely, under legitimate system operation (corresponding to the null hypothesis [44, Ch. 14]), the controller observation $Y_k$ behaves according to

$$Y_{k+1} - aY_k - U_k(Y_k) \sim \text{i.i.d. } f_W. \tag{5}$$

Note that in case of an attack, during Phase 2 ($k > L$), (5) can be rewritten as

$$V_{k+1} - aV_k - U_k(Y_k^*) = V_{k+1} - aV_k + \hat{A}V_k - \hat{A}V_k - U_k(Y_k^*) \tag{6a}$$

$$= \bar{W}_k + (\hat{A} - a)V_k, \tag{6b}$$

where (6b) follows from (4). Hence, the estimation error $(\hat{A} - a)$ dictates the ease with which an attack can be detected.

While the controller in general can carry out different statistical tests to test the validity of (5) such as the Kolmogorov–Smirnov and Anderson–Darling tests [44, Ch. 14], we consider a specific test in Sec. III-A that requires knowledge of only the second-order statistics.

### D. Deception and Detection Probabilities

Define the hijack indicator at time $k$ as

$$\Theta_k \triangleq \begin{cases} 0, & \forall j \leq k : Y_j = X_j; \\ 1, & \text{otherwise.} \end{cases} \tag{7}$$

At every time $k$, the controller uses $Y_k^*$ to construct an estimate $\hat{\Theta}_k$ of $\Theta_k$. We consider the following events.

- $\Theta_k = 0, \hat{\Theta}_k = 0$: There was no attack, and no attack was declared by the detector.
- $\Theta_k = 1, \hat{\Theta}_k = 1$: An attacker hijacked the controller observation before time $k$ but was caught by the controller/detector. In this case we say that the controller detected the attack. The detection probability at time $k$ is defined as

$$P^a_{\text{Det}} \triangleq \mathbb{P}_a \left( \hat{\Theta}_k = 1 \middle| \Theta_k = 1 \right). \tag{8}$$

- $\Theta_k = 1, \hat{\Theta}_k = 0$: An attacker hijacked the observed signal by the controller before time $k$, and the controller/detector failed to detect the attack. In this case, we say that the attacker deceived the controller or, equivalently, that the controller misdetected the attack [44, Ch. 3]. The deception probability at time $k$ is defined as

$$P^a_{\text{Dec}} \triangleq \mathbb{P}_a \left( \hat{\Theta}_k = 0 \middle| \Theta_k = 1 \right). \tag{9}$$

- $\Theta_k = 0, \hat{\Theta}_k = 1$: The controller falsely declared an attack. We refer to this event as false alarm. The false alarm probability at time $k$ is defined as

$$P^a_{\text{FA}} \triangleq \mathbb{P}_a \left( \hat{\Theta}_k = 1 \middle| \Theta_k = 0 \right). \tag{10}$$

Clearly,

$$P^a_{\text{Det}} + P^a_{\text{Dec}} = 1. \tag{11}$$

The controller wishes to achieve a low false alarm probability, while guaranteeing a low deception probability and a low control cost (5). In addition, in case of an attacker that knows (or has perfectly learned) the system gain $a$, and replaces $\{X_k\}$ of (1) with a virtual signal $\{V_k\}$ that is statistically identical and independent of it, the controller has no hope of correctly detecting the attack.

We further define the deception, detection, and false alarm probabilities w.r.t. the probability measure $\mathbb{P}$, without conditioning on $A$, and denote them by $P^k_{\text{Det}}, P^k_{\text{Dec}},$ and $P^k_{\text{FA}}$, respectively. For instance, $P^\mathbb{P}_{\text{Det}}$ is defined as

$$P^\mathbb{P}_{\text{Det}} \triangleq \mathbb{P} \left( \hat{\Theta}_k = 1 \middle| \Theta_k = 1 \right) = \int_{-\infty}^\infty P^k_{\text{Det}} f_A(a)\, da \tag{12}$$

w.r.t. a PDF $f_A$ of $A$.

### III. Statement of the results

We now describe the main results of this work. We start by describing a variance-based attack-detection test in Sec. III-A. We derive upper and lower bound on the deception probability in Sec. III-B. The proofs of the results in this section are relegated to the appendix.

#### A. Attack-Detection Variance Test

A simple and widely used test is the one that seeks anomalies in the variance, i.e., a test that examines whether the empirical variance of (3) is equal to $\mathbb{E} [W^2]$. In this way, only second-order statistics of $W$ need to be known at the controller. The price of this is of course its inability to detect higher-order anomalies.

Specifically, this test sets a confidence interval of length $2\delta > 0$ around the expected variance, i.e., it checks whether

$$\frac{1}{T} \sum_{k=1}^T \left[ Y_{k+1} - aY_k - U_k(Y_k) \right]^2 \in (\mathbb{V} [W] - \delta, \mathbb{V} [W] + \delta), \tag{13}$$

where $T$ is called the test time. That is, as is implied by (6), the attacker manages to deceive the controller ($\Theta_k = 0$) if

$$\frac{1}{T} \left( \sum_{k=1}^L W_k^2 + \sum_{k=L+1}^T (\bar{W}_k + (\hat{A} - a)V_k)^2 \right) \in (\mathbb{V} [W] - \delta, \mathbb{V} [W] + \delta).$$

Eq. (3) suggests that the false alarm probability of the variance test (13) is

$$P^a_{\text{FA}} = \mathbb{P}_a \left( \frac{1}{T} \sum_{k=1}^T W_k^2 - \mathbb{V} [W] \geq \delta \right).$$
By applying Chebyshev’s inequality and (2), we have
\[
\Pr_{\text{FA}}^T \leq \min \left(1, \frac{\Var[W^2]}{\delta^2 T} \right) = \min \left(1, \frac{\E_p [W^4] - \E_p^2 [W^2]}{\delta^2 T} \right).
\]
(14)

As a result, as \( T \to \infty \) the probability of false alarm tends to zero. Hence, in this limit, we are left with the task of determining the behavior of the deception probability \( \Pr \). We note that the asymptotic assumption \( T \to \infty \) simplifies the presentation of the results. Nonetheless, a similar treatment can be obtained in the non-asymptotic case.

**B. Bounds on the Deception Probability Under the Variance Test**

In what follows, we assume that there exists \( \beta > 0 \), such that
\[
\lim_{T \to \infty} \frac{1}{T} \sum_{k=L+1}^{T} V_k^2 = 1/\beta \quad \text{a.s. w.r.t. } \mathbb{P}_a.
\]
(15)

Remark 2. Assuming the control policy is memoryless, namely \( U_k = o(T) \), the process \( V_k \) is Markov for \( k \geq L + 1 \). By further assuming that \( L = o(T) \) and using the generalization of the law of large numbers for Markov processes [45], we deduce
\[
\lim_{T \to \infty} \frac{1}{T} \sum_{k=L+1}^{T} V_k^2 \geq \Var[W] \quad \text{a.s. w.r.t. } \mathbb{P}_a.
\]

Consequently, in this case we have \( \beta \leq 1/\Var[W] \). In addition, when the control policy is linear and stabilizes \( \hat{A} \), that is \( U_k = -\Omega Y_k \) and \( |\hat{A} + \Omega| < 1 \), it is easy to verify that \( (\ref{eq:var}) \) holds true for \( \beta = (1 - (\hat{A} + \Omega)^2)/\Var[W] \).

We now provide lower and upper bounds on the deception probability \( \Pr_{\text{Dec}}^T \) of any linear two-phase AIA [4] where \( \hat{A} \) in [4] is constructed using any arbitrary learning algorithm.

1) **Lower Bound:** To provide a lower bound on the deception probability \( \Pr_{\text{Dec}}^T \), we consider a specific estimate of \( \hat{A} \) at the conclusion of the first phase by the attacker, assuming a controller that uses the variance test (13). To this end, we use least-squares (LS) estimation due to its efficiency and amenability to recursive update over observed incremental data, which makes it the method of choice for many applications of real-time parametric identification of dynamical systems [40], [46]–[52].

The LS algorithm approximates the overdetermined system of equations
\[
\begin{pmatrix}
X_1 \\
X_2 \\
\vdots \\
X_{L-1} \\
X_L
\end{pmatrix} = \begin{pmatrix}
U_1 \\
U_2 \\
\vdots \\
U_{L-1} \\
U_L
\end{pmatrix},
\]
by minimizing the Euclidean distance
\[
\hat{A} = \arg\min_A \|AX_k + U_k\|,
\]
to estimate (or “identify”) the plant, the solution to which is
\[
\hat{A} = \frac{\sum_{k=1}^{L} -U_k X_k}{\sum_{k=1}^{L} X_k^2} \quad \text{a.s. w.r.t. } \mathbb{P}_a
\]
(16)

Remark 3. Since we assumed \( W_k \) for all time \( k \) has a PDF, the probability that \( X_k = 0 \) is zero. Consequently, (16) is well-defined.

Using LS estimation (16) achieves the following asymptotic deception probability.

**Theorem 1.** Consider any linear two-phase AIA [4] with fictitious-sensor reading power that satisfies [15] and a control policy \( \{U_k\} \). Then, the asymptotic deception probability when using the variance test (13) is bounded from below as
\[
\lim_{T \to \infty} \Pr_{\text{Dec}}^T = \mathbb{P}_a \left( |\hat{A} - a| < \sqrt{\delta \beta} \right) \geq \mathbb{P}_a \left( \sum_{k=1}^{L} (aX_k + U_k)^2 < \delta \beta \right).
\]

Remark 4. Thm. 1 guarantees \( \lim_{T \to \infty} \Pr_{\text{Dec}}^T = 1 \) for the choice \( U_k = -aX_k \), where \( Y_k = X_k \) during the exploration phase. \( U_k = -aX_k \) is the optimal control policy for LQ control [3] for \( r = 0 \) (no penalty on the control actions). An important consequence is that, for this choice, even without having any prior knowledge of the open-loop gain of the plant, the attacker can still carry out a successful attack.

2) **Upper Bound:** We derive an upper bound on the deception probability for the case of a uniformly distributed \( A \) over a symmetric interval \([-R, R]\). We assume the attacker knows the distribution of \( A \) (including the value of \( R \)), whereas the controller knows the true value of \( A \) (as before). Similar results can be obtained for other interval choices.

**Theorem 2.** Let \( A \) be distributed uniformly over \([-R, R]\) for some \( R > 0 \), and consider any control policy \( \{U_k\} \) and any linear two-phase AIA [4] with fictitious-sensor reading power (15) that satisfies \( \sqrt{\beta} \leq R \). Then, the asymptotic deception probability when using the variance test (13) is bounded from above as
\[
\lim_{T \to \infty} \Pr_{\text{Dec}}^T = \lim_{T \to \infty} \mathbb{P} \left( \hat{A} = 0 \mid \Theta_T = 1 \right) = 1 - \frac{1}{2R} \int_{-R}^{R} \mathbb{P}_a \left( |\hat{A} - a| \geq \sqrt{\delta \beta} \right) da.
\]
(17a)

\[
\leq \Lambda \triangleq I(A; Z_1^T) + 1 \log \left( \frac{R}{\sqrt{\delta \beta}} \right).
\]
(17b)

\[
\Lambda \leq \sum_{k=1}^{L} D \left( \mathbb{P}_{X_k \mid Z_1^{k-1}}^k, \mathbb{P}_{Q_{X_k \mid Z_1^{k-1}}} \right) + 1 \log \left( \frac{R}{\sqrt{\delta \beta}} \right).
\]
(17c)

In addition, if for all \( k \in \{1, \ldots, L\} \) \( A \to (X_k, Z_1^{k-1}) \to U_k \) is a Markov chain, then for any sequence of probability measures \( \{Q_{X_k \mid Z_1^{k-1}}\} \), such that for all \( k \in \{1, \ldots, L\} \)
\[
\mathbb{P}_{X_k \mid Z_1^{k-1}} \ll Q_{X_k \mid Z_1^{k-1}}, \text{ we have}
\]
\[
\Lambda \leq \sum_{k=1}^{L} D \left( \mathbb{P}_{X_k \mid Z_1^{k-1}, A}^{k}, Q_{X_k \mid Z_1^{k-1}} \right) + 1 \log \left( \frac{R}{\sqrt{\delta \beta}} \right).
\]

Remark 5. By looking at the numerator in (17c), it follows that the bound on the deception probability becomes looser as the amount of information revealed about \( \hat{A} \) by the observation \( Z_1^L \) increases. On the other hand, by looking at the denominator,
the bound becomes tighter as $R$ increases. This is consistent with the observation of Zames [53] (see also [49]) that SysID becomes harder as the uncertainty about the open-loop gain of the plant increases; in our case, a larger uncertainty interval $R$ corresponds to a worse estimation of the open-loop gain $A$ by the attacker, which leads, in turn, to a decrease in the attacker achievable deception probability. The denominator can also be interpreted as the intrinsic uncertainty of $A$ when this is observed at resolution $\sqrt{\delta^2}$, as it corresponds to the entropy of the random variable $A$ when it is quantized at such resolution.

In conclusion, Thm. 2 provides two upper bounds on the deception probability. The first (17) clearly shows that increasing the privacy of the open-loop gain $A$—manifested in the mutual information between $A$ and the state-and-control trajectory $Z^k_T$ during the exploration phase—reduces the deception probability. The second bound (18) allows freedom in choosing the auxiliary probability measure $Q_{X_k|Z^{k-1}_k}$, making it a rather useful bound. An important instance is that of an i.i.d. Gaussian plant disturbance sequence $W_k \sim \mathcal{N}(0, \sigma^2)$; by choosing $Q_{X_k|Z^{k-1}_k} \sim \mathcal{N}(0, \sigma^2)$, for this case, for all $k \in \mathbb{N}$, we can rewrite the upper bound (18) in term of $E_\phi \left[(AX_k-1 + U_k-1)^2\right]$ as follows.

Corollary 1. Under the assumptions of Thm. 2 if for all $k \in \{1, \ldots, L\} \ A \rightarrow (X_k, Z^{k-1}_k) \rightarrow U_k$ is a Markov chain, and $W_k \sim \mathcal{N}(0, \sigma^2)$ is an i.i.d. Gaussian plant disturbance sequence, the following upper bound on the asymptotic deception probability holds:

$$\lim_{T \rightarrow \infty} P^T_{\text{Dec}} \leq G(Z^T_1),$$

where

$$G(Z^T_1) \triangleq \frac{\log_2 \sum_{k=1}^L E_\phi \left[(AX_k-1 + U_k-1)^2\right] + 1}{\log_2 \left(R \sqrt{\delta^2}\right)}. \quad (20)$$

Remark 6. While the upper bound in (17) is valid for all control policies, the upper bound in (18), and consequently also the one in (19), is only valid for control policies where $A \rightarrow (X_k, Z^{k-1}_k) \rightarrow U_k$ form a Markov chain for all $k \in \{1, \ldots, L\}$. To demonstrate this, choose $U_k = -AX_k$ and evaluate the bounds in (17) and (19). Clearly (20) is finite. On the other hand $I(A; Z^{T}_1)$ and hence also the upper bound in (17), is infinite, since, given $X_k$ and $U_k$, $A$ can be fully determined.

C. Watermarking

To increase the security of the system, at any time $k$ the controller can add an authentication (or watermarking) signal $\Gamma_k$ to an unauthenticated control policy $\{\hat{U}_k|k \in \mathbb{N}\}$:

$$U_k = \hat{U}_k + \Gamma_k, \quad k \in \mathbb{N}. \quad (21)$$

We refer to such a control policy $U_k$ as the authenticated control policy $\hat{U}_k$. We denote the states of the system that would be generated if only the unauthenticated control signal $\hat{U}_k$ were applied by $\hat{X}_k$, and the resulting trajectory—by $\hat{Z}_k \triangleq (\hat{X}_k, \hat{U}_k)$. A “good” authentication signal entails little increase in the control cost (3) compared to its unauthenticated version while providing enhanced detection probability (9) and/or false alarm probability.

Remark 7. In both the replay-attack [32] and the statistical-duplicate [38] models no access to the control signal $U_k$ by the attacker was allowed. Thus, to improve the detection probability of the controller in case of an attack, one could add an authentication/watermarking signal, which helped the controller to identify abnormalities by correlating the input watermarking signal with its contribution to the sensor reading.

However, since in the statistical-duplicate setting full system knowledge at the attacker was assumed, if the attacker has the access to the control signal it could easily simulate the contribution of the any inscribed watermarking signal to the sequence of fictitious sensor readings. In contrast, in the replay-attack setting, no system knowledge is assumed, rendering any knowledge of the control signal useless, unless learning the plant dynamics is invoked. In our setup the attacker has full access to the control signal. However, in contrast to the statistical-duplicate setting, it cannot perfectly simulate the effect of the control signal as it lacks knowledge of the open-loop gain. Thus, the watermarking signal here is used for a different purpose—to impede the learning process of the attacker.

At first glance, one may envisage that superimposing any watermarking signal $\Gamma_k$ on top of the control policy $\{\hat{U}_k|k \in \mathbb{N}\}$ would necessarily enhance the detectability of an attack, since the observations of the attacker are in this case noisier. However, it turns out that injecting a strong noise may in fact speed up the learning process as it improves the power of the signal magnified by the open-loop gains with respect to the observed noise [54].

The following corollary proposes a class of watermarking signals that provide enhanced guarantees on the deception probability $P^T_{\text{Dec}}$.

Corollary 2. For any control policy $\{\hat{U}_k|k \in \mathbb{N}\}$ with trajectory $\hat{Z}^T_1 = (\hat{X}^T_1, \hat{U}^T_1)$ and its corresponding authenticated control policy $U^T_1$ (21) with trajectory $Z^T_1 = (X^T_1, U^T_1)$, under the assumptions of Corollary 7 if for all $k \in \{1, \ldots, L\}$

$$E_\phi \left[\Psi^2_{k-1} + 2\Psi_{k-1}(AX_{k-1} + U_{k-1})\right] < 0, \quad (22)$$

then letting $\Psi_{k-1} \triangleq \sum_{j=1}^{k-1} A^{k-1-j} \Gamma_j$, the following majorization holds:

$$G(Z^T_1) < G(\hat{Z}^T_1),$$

where $G$ is defined in (20).

IV. SIMULATIONS

In this section, we compare the empirical performance of the variance-test with our developed bounds. At every time $T$, the controller tests the empirical variance for abnormalities over a detection window $[T-T^\prime+1, T]$, where $T^\prime \leq T$, using a confidence interval $2\delta > 0$ around the expected variance (13). When $T^\prime = T$ the statistical test used in the simulation, the hijack indicator $\Theta_T$, and its estimate $\hat{\Theta}_T$ at the controller reduce to the definitions of the variance test in (13), the hijack indicator in (7), and the estimate of the latter of Sec. II-D respectively. In the interest of brevity, we state the exact simulation parameters only in the figure captions.
A. False Alarm Rate

We start by examining the false alarm rate $P_{FA}^{\alpha,T}$ under the variance test as a function of the detection window size $T'$. We depict, in Fig. 2, the empirical false alarm rate evaluated using the Monte Carlo method along with the theoretical bound of (14). Both the theoretical bound and the empirical performance present a similar behavior: the false alarm rate is large for small values of the detection window $T'$, and decays to zero for $T' \to \infty$. We note that since Chebyshev’s inequality is not tight in general there exists a gap between the empirical and theoretical performance that can be closed using exponential bounds that require further statistical knowledge of $W_k$.

B. Replay Attack

The variance-test detection rate $P_{FA}^{\alpha,T}$ of a replay attack as a function of the detection window size $T'$ is depicted in Fig. 3 along with the corresponding false alarm rate. No watermarking signals are used in the simulation. The figure clearly shows that the variance-test detection rate of a replay attack saturates when the detection window size $T'$ goes to infinity for a fixed confidence interval $2\delta$. On the other hand, as discussed in Sec. IV-A the false alarm rate decays to zero as $T' \to \infty$, in this case.

Remark 8. Instead of fixing the confidence interval $2\delta$, one may choose to hold the false alarm rate fixed to a prescribed value by appropriately adjusting the confidence interval. Under this setup, the detection rate will decrease to zero upon increasing the detection window size.

C. Adaptive Integrity Attack

Fig. 4 compares the attacker’s success rates $P_{Dec}^{\alpha,T}$, as a function of the detection window size $T'$, of the replay attack and the AIA.

As discussed in Sec. IV-A, the false alarm rate decays to zero as the size of the detection window $T'$ tends to infinity. Hence, for a sufficiently large detection window size, the attacker’s success rate could potentially tend to one.

Indeed, such a behavior is observed in Fig. 4, where the controller executes a control policy under the LQ cost criterion (2), with $q = 1$ and $r = 0$ (no penalty on the control action): the success rate of the AIA tends to one as $T' \to \infty$, whereas the success rate of the replay attack saturates at a lower value.

Fig. 5 depicts a similar comparison for $U_k = -0.77aY_k$, which corresponds to a control policy under the LQ cost criterion (3) with $r = 1$ and $q = 2.239$. In this case, the success rate of the AIA is lower than that for $r = 0$ (depicted in Fig. 4). This can also be seen analytically by noting that the attacker uses the LS algorithm (16) of Sec. III-B.1 to construct $\hat{A}$, for which the deception probability of the attacker as $T \to \infty$ (characterized in Thm. 1) is the one for the optimal control policy for $r = 0$ ($U_k = -aY_k$).

Fig. 6 compares the attacker’s success rate $P_{Dec}^{\alpha,T}$ as a function of the exploration phase duration $L$ of four different control policies: I) $U_k = -aY_k$, II) $U_k = -aY_k + 0.5$, III) $U_k = -0.9aY_k + 0.5$, IV) $U_k = -0.8aY_k + 0.4$. As is illustrated in Fig. 5 and in agreement with Thm. 1 the deception probability decreases as the control policy gets farther away from the optimal control policy under an LQ control cost (5) and $r = 0$, when the attacker uses the LS algorithm (16) to construct $\hat{A}$.

D. Watermarking

We now evaluate the detection rate as a function of the power of the watermarking signal $\Gamma_k$ under a variance test and an AIA attack that uses the LS algorithm (16) to construct $\hat{A}$. To this end, we fix the detection window to be $T = T' = 800$, which guarantees a negligible false alarm probability as discussed in Sec. IV-A and use Gaussian i.i.d. zero-mean watermark signals $\{\Gamma_k\}$ as in (21), with different powers; the control action is equal to $U_k = -0.85aY_k + \Gamma_k$. The performance is depicted in Fig. 6.
exploitation phases are separate from each other. However, when in Fig. 7. The attacker’s success rate $P^{1,2,800}_{\text{Dec}} = 1 - P^{1,2,800}_{\text{Det}}$ versus the duration of the exploration phase $L$. The system parameters are $\delta = 0.1$, $T = 800$, $a = 1.2$, and $\{W_k\}$ are i.i.d. standard Gaussian. The attacker used the LS algorithm (16) to construct $A$. The four curves correspond to four different control policies; case I: $U_k = -aY_k$, case II: $U_k = -aY_k + 0.5$, case III: $U_k = -0.9aY_k + 0.5$, case IV: $U_k = -0.8aY_k + 0.4$. A semi-logarithmic scale for the duration of the exploration phase $L$ is used. 500 Monte Carlo simulations were performed.

Fig. 5: The attacker’s success rate $P^{1,2,800}_{\text{Dec}} = 1 - P^{1,2,800}_{\text{Det}}$ versus the duration of the exploration phase $L$. The system parameters are $\delta = 0.1$, $T = T' = 800$, $a = 1.2$, and $\{W_k\}$ are i.i.d. standard Gaussian. The attacker used the LS algorithm (16) to construct $A$. The four curves correspond to four different control policies; case I: $U_k = -aY_k$, case II: $U_k = -aY_k + 0.5$, case III: $U_k = -0.9aY_k + 0.5$, case IV: $U_k = -0.8aY_k + 0.4$. A semi-logarithmic scale for the duration of the exploration phase $L$ is used. 500 Monte Carlo simulations were performed.

**E. Random Open-Loop Gain**

Thm. [1] provides a lower bound on the deception probability conditioned on $A = a$. Hence, by applying the law of total probability w.r.t. the PDF $f_A$ of $A$ as in (12), we can apply the result of Thm. [1] to provide a lower bound also on the average deception probability for a random $A$. In this context, Fig. [7] compares the lower and upper bounds on the deception probability provided by Thm. [1] and Corol. [1], respectively, where $A$ is distributed uniformly over $[-1, 1]$. Thm. [19] is valid when the control input is not a function of random variable $A$; hence, we assumed $U_k = -0.045Y_k$. In addition, a simple calculation shows that if the control input is proportional to $Y_k$, then the lower bound provided in Thm. [1] is not a function of the duration of the exploration phase $L$, as is demonstrated in Fig. [7].

**V. DISCUSSION AND FUTURE WORK**

**A. Online Learning**

In our analysis we assumed that the exploration and exploitation phases are separate from each other. However, when the attacker hijacks the input and output control signals, it can continue “exploring” the plant (see Fig. [15] and improving its estimation/identification of the plant. Securing systems against such online attacks is an interesting research problem.

**B. Deception–Cost Tradeoff**

Denote by $\hat{U}_k$ the optimal unauthenticated control signal (recall Sec. III-C) and its corresponding LQ cost—by $J_T^*$. We further define an authenticated control signal $U_k = f_k(U_k)$ at time $k$, where $f_k$ is some possibly random function; we define the corresponding cost by $J_T$ and the increase in cost due to the authentication, compared to the optimal (unauthenticated) control cost $J_T^*$—by $\Delta J_T := J_T - J_T^*$. By invoking Thm. [2], we can optimize the upper bound (17) on the deception probability by solving the optimization problem

$$\min_{\forall f_k: \Delta J_T \leq \theta} I(A; X^T(U^T_I(U^T_2)), U^T_T))$$

**C. Oblivious Controller**

A more realistic scenario is the one in which neither the attacker nor the controller are aware of the open-loop gain of the plant. In this scenario, both parties strive to learn the open-loop gain—posing two conflicting objectives to the controller, who on the one hand, wishes to speed up the learning process to itself, but on the other hand wishes to slow down the learning process of the eavesdropping attacker.

In such a situation, standard adaptive control methods [40], [47]–[49] are clearly insufficient, as no asymmetry between the attacker and the controller can be achieved under the setting of Sec. II. To create a security leverage over the attacker, the controller needs to utilize watermarking. Interestingly, a properly designed watermarking signal would enjoy a positive double effect by facilitating the learning process of the controller while hindering that of the attacker at the same time.
LS algorithm (16) to construct $\hat{A}$ performed. T equals $\bar{L}$ with an exploration phase of length for all $k \leq 800$, and $\{W_k\}$ are i.i.d. Gaussian with zero mean and a variance of 0.01. 10000 Monte Carlo simulations were performed.

Fig. 7: Comparison of the lower and upper bounds on the deception probability, of Thm. 1 and Corol. 1, respectively, where $A$ is distributed uniformly over $[-1,1]$. The controller uses a variance test whereas the attacker carries out an AIA. The system parameters are $\delta = 0.1$, $a = 1.2$, $U_k = -0.045Y_k$ for all $1 \leq k \leq 800$, and $\{W_k\}$ are i.i.d. Gaussian with zero mean and a variance of 0.01. 10000 Monte Carlo simulations were performed.

Finally, we note that unless the controller is able to detect an MITM attack (the attacker’s exploitation phase), its learning process will be hampered by the fictitious signal that is generated according to the virtual system of the attacker 4.

D. Vector Systems

We concentrated on scalar systems in this work. The extension of the established results to the vector (possibly partially-observable) case is an interesting research venue we would like to explore.

E. Continuous Testing

Throughout this work, we have assumed that the controller tests the integrity of the system at a specific time step $T$, that is taken, in turn, to be very large (tends to infinity). Since the controller does not know the exact time instant at which an attack might occur, a more realistic scenario would be that of continuous testing, i.e., that in which the integrity of the system is tested at every time step and where the false alarm and deception probabilities are defined with a union over time. We leave this treatment for future research.

F. A Unified View of Cyber-Physical Security

A unified view of cyber-physical systems security is presented in Fig. 8. The different ways that switches $\phi_1, \phi_2, \phi_3$ and $\phi_4, \phi_5, \phi_6$ in Fig. 8 can be connected together correspond to different attack scenarios, as detailed below.

- **Denial-of-service (DoS) attack** [23]–[26]. In this attack the attacker disrupts the service by overloading the system with superfluous requests. This can be modeled by an attacker that opens and closes switches $\phi_1 \leftrightarrow \phi_2$ or $\phi_4 \leftrightarrow \phi_5$, interchangeably, either in a random or a deterministic fashion.

- **False data-injection attack** [18], [19]. In this attack, a false input signal is injected either to the plant or to the controller, corresponding to switches $\phi_2 \leftrightarrow \phi_3$ and $\phi_4 \leftrightarrow \phi_5$, or switches $\phi_1 \leftrightarrow \phi_2$ and $\phi_5 \leftrightarrow \phi_6$ being closed, respectively.

- **Man-in-the-middle (MITM) attack.** In this model the attacker intervenes in two places at the control loop, corresponding to switches $\phi_2 \leftrightarrow \phi_3$ and $\phi_5 \leftrightarrow \phi_6$ being closed. This scenario reduces to the one depicted in Fig. 1b and subsumes the replay attack, the statistical-duplicate attack, and the AIA discussed in this work.

---

**APPENDIX I**

**PROOF OF THEOREM 1**

We break the proof of Thm. 1 into several lemmas that are stated and proved next.

In the case of linear two-phase attacks, in the limit of $T \to \infty$, the empirical variance, which is used in the variance
test \([13]\), can be expressed in terms of the estimation error of the open-loop gain as follows.

**Lemma 1.** Consider any linear two-phase attack \([4]\) with fictitious-sensor reading power that satisfies \([15]\) and some control policy \([U_k]\). Then, the variance test \([13]\) reduces to

\[
\lim_{T \to \infty} \frac{1}{T} \sum_{k=1}^{T} (Y_{k+1} - a Y_k - U_k(Y_1^k))^2 = \frac{\text{Var}[W]}{\beta} + \frac{(\hat{A} - a)^2}{\beta} \quad \text{a.s. w.r.t. } \mathbb{P}_a.
\]

**Proof.** Since the exploration phase of a linear two-phase attack starts at time \(k = L + 1\), using \([1]\) and \([6]\) we have

\[
\begin{align*}
&1 \sum_{k=1}^{T} (Y_{k+1} - a Y_k - U_k(Y_1^k))^2 \\
&= \frac{1}{T} \left( \sum_{k=1}^{L} W_k^2 + \sum_{k=L+1}^{T} (\hat{W}_k + (\hat{A} - a)V_k)^2 \right) \\
&= \frac{1}{T} \left( \sum_{k=1}^{L} W_k^2 + \sum_{k=L+1}^{T} \hat{W}_k^2 \right) \\
&+ \frac{(\hat{A} - a)^2}{T} \sum_{k=L+1}^{T} V_k^2 + \frac{2(\hat{A} - a)}{T} \sum_{k=L+1}^{T} \hat{W}_k V_k .
\end{align*}
\]

Using \([48, \text{Lemma 2, part iii}]\) the last term in \((23b)\) reduces to

\[
\sum_{k=L+1}^{T} \hat{W}_k V_k = o \left( \sum_{k=L+1}^{T} V_k^2 \right) + O(1) \quad \text{a.s.}
\]

in the limit \(T \to \infty\).

Note further that

\[
\lim_{T \to \infty} \frac{1}{T} \left( \sum_{k=1}^{L} W_k^2 + \sum_{k=L+1}^{T} \hat{W}_k^2 \right) = \text{Var}[W] \quad \text{a.s. (24)}
\]

by the strong law of large numbers \([55]\).

Substituting \([15], (23)\), and \((24)\) in \((23)\) and taking \(T\) to infinity concludes the proof of the lemma. \(\square\)

**Lemma 2.** Consider any linear two-phase attack \([4]\) with fictitious-sensor reading power that satisfies \([15]\) and a control policy \([U_k]\). Then, the asymptotic deception probability, under the variance test \([13]\), is equal to

\[
\lim_{T \to \infty} P^{a,T}_{\text{Dec}} = \mathbb{P}_a \left( |\hat{A} - a| < \sqrt{\delta \beta} \right).
\]

**Proof.** Under the variance test, \(\text{Var}[W] + \frac{(\hat{A} - a)^2}{\beta} \in (\text{Var}[W] - \delta, \text{Var}[W] + \delta)\) and zero otherwise. Consequently,

\[
\lim_{T \to \infty} P^{a,T}_{\text{Dec}} = \mathbb{P}_a \left( (\hat{A} - a)^2 \in (-\delta \beta, \delta \beta) \right),
\]

and the result follows. \(\square\)

The following lemma establishes an upper bound on the squared estimation error \((\hat{A} - a)^2\) in terms of the state and disturbance sequences of plant, when \(\hat{A}\) is constructed using LS estimation \([16]\).

**Lemma 3.** If the attacker constructs \(\hat{A}\) using LS estimation \([16]\), then

\[
(\hat{A} - a)^2 \leq \frac{\sum_{k=1}^{L} (X_{k+1} - W_k)^2}{\sum_{k=1}^{L} X_k^2} \quad \text{a.s. w.r.t. } \mathbb{P}_a
\]

**Proof.** This proof essentially follows from \([49, \text{Lemma 1}]\) for our scalar plant \([1]\). As mentioned before, since \(W_k\) for all time \(k\) has a PDF the probability that \(X_k = 0\) is zero. Hence, the random variables \(S_k = U_k/X_k\) are well-defined except on a measure zero set for all \(k\). Thus, we have

\[
a - \hat{A} = \frac{\sum_{k=1}^{L} (a - S_k)X_k^2}{\sum_{k=1}^{L} X_k^2} \quad \text{a.s. w.r.t. } \mathbb{P}_a
\]

By further applying the Cauchy–Schwarz inequality, we deduce

\[
(a - \hat{A})^2 \leq \frac{\sum_{k=1}^{L} ((a - S_k)X_k)^2}{\sum_{k=1}^{L} X_k^2} \quad \text{a.s. (25)}
\]

Finally, by noting that \((a - S_k)X_k = aX_k - U_k = X_{k+1} - W_k\) and substituting it in \((25)\) concludes the proof. \(\square\)

The proof of Thm. \([2]\) now follows by combining the results of Lemmata \([2]\) and \([3]\) and noting that \(X_{k+1} - W_k = aX_k + U_k\).

**APPENDIX II**

**PROOF OF THEOREM [2]**

The proof is based on the continuous Fano inequalities [56]. We start by proving \([17]\). Since the attacker observed the plant state and control input during the exploration phase which lasts \(L\) time steps, and since \(A \to (X_1^L, U_1^L) \to A\) constitutes a Markov chain, using the continuous domain version of Fano’s inequality [56, Prop. 2], we have

\[
\inf_A \mathbb{P} \left( |A - \hat{A}| \geq \sqrt{\delta \beta} \right) \geq 1 - \frac{I(A; Z^T_1) + 1}{\log \frac{R}{\sqrt{\delta \beta}}},
\]

whenever \(\sqrt{\delta \beta} \leq R\).

Using \([11]\) and Lem. \([2]\) we deduce

\[
\lim_{T \to \infty} P^{a,T}_{\text{Det}} = \mathbb{P}_a \left( |\hat{A} - a| \geq \sqrt{\delta \beta} \right).
\]

Consequently, using \([12]\) and the dominated convergence theorem [55] it follows that

\[
\lim_{T \to \infty} \mathbb{P} \left( \Theta_T = 1 \mid \Theta_T = 1 \right) = \frac{1}{2R} \int_{-R}^{R} \mathbb{P}_a \left( |\hat{A} - a| \geq \sqrt{\delta \beta} \right) da = \mathbb{P} \left( |A - \hat{A}| \geq \sqrt{\delta \beta} \right).
\]
By further applying (26), we obtain
\[
\lim_{T \to \infty} \mathbb{P} \left( \Theta_T = 1 \mid \Theta_T = 1 \right) \geq 1 - \frac{I(A; Z_T^1) + 1}{\log(R/\sqrt{\delta \beta})},
\]
which proves (17) by recalling (11).

To prove (18), we further bound \( I(A; Z_T^1) \) from above via KL divergence manipulations; similar arguments have been previously used, e.g., in [57]. The proof of the following lemma follows the arguments of [49], and is detailed here for completeness.

Lemma 4. Assume the following Markov chain \( A \to (X_k, Z_{1:k}^1) \to U_k \) all \( k \in \{1, \ldots, L\} \). Let \( \{Q_{X_k}|Z_{1:k}^1\} \) be a sequence of probability measures satisfying \( \mathbb{P}_{X_k|Z_{1:k}^1} \ll Q_{X_k}|Z_{1:k}^1 \) for all \( k \). Then, for all \( k \), we have
\[
I(A; Z_T^k) = \sum_{k=1}^{L} I(A; X_k|Z_{1:k}^1) \leq \sum_{k=1}^{L} D \left( \mathbb{P}_{X_k|Z_{1:k}^1} \mid \mathbb{P}_{Z_{1:k}^1} \right).
\]

Proof. We start by applying the chain rule for mutual informations to \( I(A; Z_T^k) \) as follows.
\[
I(A; Z_T^k) = \sum_{k=1}^{L} I(A; X_k|Z_{1:k}^1).
\]

We next bound \( I(A; Z_k|Z_{1:k}^1) \) from above:
\[
I(A; Z_k|Z_{1:k}^1) = I(A; X_k, U_k|Z_{1:k}^1) = D \left( \mathbb{P}_{X_k|Z_{1:k}^1} \mid \mathbb{P}_{Z_{1:k}^1} \right) = \mathbb{E}_\mathbb{P} \log \frac{d\mathbb{P}_{X_k|Z_{1:k}^1}}{d\mathbb{P}_{Z_{1:k}^1}} - \mathbb{E}_\mathbb{P} \log \frac{d\mathbb{P}_{U_k|Z_{1:k}^1}}{d\mathbb{P}_{Z_{1:k}^1}} \leq D \left( \mathbb{P}_{X_k|Z_{1:k}^1} \mid \mathbb{P}_{Z_{1:k}^1} \right).
\]

where we substitute the definition of \( Z_k \equiv (X_k, U_k) \) to arrive at (28a). Using the chain rule for mutual informations and the Markovity assumption \( A \to (X_k, Z_{1:k}^1) \to U_k \), we use the definition of the conditional mutual information in terms of the conditional KL divergence (recall Sec. I-A) to attain (28c), and (28d) follows from the non-negativity property of the KL divergence.

Substituting (28d) in (27) concludes the proof.

APPENDIX III
PROOF OF COROLLARY I

Set \( Q_{X_k}|Z_{1:k}^1 \sim \mathcal{N}(0, \sigma^2) \). Then, \( \mathbb{P}_{X_k|Z_{1:k}^1} = \mathcal{N}(AX_k-1 + U_{k-1}, \sigma^2) \), and hence the measure domination condition \( \mathbb{P}_{X_k|Z_{1:k}^1} \ll Q_{X_k}|Z_{1:k}^1 \) holds.

\[
D(\mathbb{P}_{X_k|Z_{1:k}^1}, \mathbb{P}_{Z_{1:k}^1}) = \mathbb{E}_\mathbb{P} \left[ D \left( \mathcal{N}(AX_k-1 + U_{k-1}, \sigma^2) \mid \mathcal{N}(0, \sigma^2) \right) \right] \]
\[
= \log \frac{1}{2\sigma^2} \mathbb{E}_\mathbb{P} \left[ (AX_k-1 + U_{k-1})^2 \right].
\]

The result follows by combining (13) and (29).

APPENDIX IV
PROOF OF COROLLARY 2

Using (1) and (21), we can rewrite \( \tilde{X}_k \) and \( X_k \) explicitly as follows:
\[
\tilde{X}_k = A^k X_0 + \sum_{j=1}^{k-1} A^{k-1-j}(U_j + W_j),
\]
\[
X_k = A^k X_0 + \sum_{j=1}^{k-1} A^{k-1-j}(U_j + W_j) = A^k X_0 + \sum_{j=1}^{k-1} A^{k-1-j}(U_j + \Gamma_j + W_j) = \tilde{X}_k + \Psi_{k-1}.
\]

Thus, the following relation holds:
\[
AX_k-1 + U_{k-1} = AX_{k-1} + \tilde{U}_{k-1} + \Psi_{k-1}.
\]

By comparing
\[
G(Z_L^1) = \frac{\log e \sum_{k=1}^{L} \mathbb{E}_\mathbb{P} \left[ (AX_{k-1} + \tilde{U}_{k-1} + \Psi_{k-1})^2 \right] + 1}{\log(R/\sqrt{\delta \beta})}
\]
with
\[
G(Z_L^1) = \frac{\log e \sum_{k=1}^{L} \mathbb{E}_\mathbb{P} \left[ (AX_{k-1} + \tilde{U}_{k-1} + \Psi_{k-1})^2 \right] + 1}{\log(R/\sqrt{\delta \beta})},
\]
in which we have utilized (30), and provided (22), we arrive at \( G(Z_L^1) > G(Z_L^1) \).

REFERENCES

[1] K.-D. Kim and P. R. Kumar, “Cyber–physical systems: A perspective at the centennial,” Proceedings of the IEEE, vol. 100, no. Special Centennial Issue, pp. 1287–1308, 2012.
[2] A. A. Cardenas, S. Amin, and S. Sarathy, “Secure control: Towards survivable cyber-physical systems,” System, vol. 1, no. a2, p. a3, 2008.
[3] H. Sandberg, S. Amin, and K. H. Johansson, “Cyberphysical security in networked control systems: An introduction to the issue,” IEEE Control Systems, vol. 35, no. 1, pp. 20–23, 2015.
[4] Y. Mo, T. H. J. Kim, K. Brancik, D. Dickinson, H. Lee, A. Perrig, and B. Sinopoli, “Cyber–physical security of a smart grid infrastructure,” Proceedings of the IEEE, vol. 100, no. 1, pp. 195–209, 2012.
[5] Y. Y. Haines, “Risk of terrorism to cyber-physical and organizational-societal infrastructures,” Public Works Management & Policy, vol. 6, no. 4, pp. 231–240, 2002.
[6] A. S. Brown, “SCADA vs. the hackers,” Mechanical Engineering, vol. 124, no. 12, p. 37, 2002.
[7] S. Checkoway, D. McCoy, B. Kantor, D. Anderson, H. Shacham, S. Savage, K. Koscher, A. Czeskis, F. Roessler, T. Kohno, et al., “Comprehensive experimental analyses of automotive attack surfaces,” in USENIX Security Symposium. San Francisco, 2011, pp. 77–92.
[8] K. Koscher, A. Czeskis, F. Roessler, S. Patel, T. Kohno, S. Checkoway, D. McCoy, B. Kantor, D. Anderson, H. Shacham, et al., “Experimental security analysis of a modern automobile,” in Security and Privacy (SP), 2010 IEEE Symposium on. IEEE, 2010, pp. 447–462.
[9] C. Miller and C. Valasek, “Remote exploitation of an unaltered passenger vehicle,” Black Hat USA, vol. 2015, p. 91, 2015.

[10] J. Slay and M. Miller, “Lessons learned from the maroochy water breach,” in International Conference on Critical Infrastructure Protection. Springer, 2007, pp. 73–82.

[11] G. Liang, S. R. WELLER, J. Zhao, F. Luo, and Z. Y. Dong, “The 2015 ukrainian blackout: Implications for false data injection attacks,” IEEE Transactions on Power Systems, vol. 32, no. 4, pp. 3317–3318, 2017.

[12] R. M. Lee, M. J. Assante, and T. Conway, “German steel mill cyber attack,” Industrial Control Systems, vol. 30, p. 62, 2014.

[13] D. P. Fidler, “Was stuxnet an act of war? decoding a cyberattack,” IEEE Security & Privacy, vol. 9, no. 4, pp. 56–59, 2011.

[14] R. Langner, “Stuxnet: Dissecting a cyber warfare weapon,” IEEE Security & Privacy, vol. 9, no. 3, pp. 49–51, 2011.

[15] T. Chen and S. Abu-Nimeh, “Lessons from stuxnet,” Computer, vol. 44, no. 4, pp. 91–93, 2011.

[16] N. Falliere, L. O. Murchu, and E. Chien, “W32. stuxnet dossier,” 2011.

[17] G. McDonald, L. O. Murchu, S. Doherty, and E. Chien, “Stuxnet 0.5: The missing link,” Symantec Report, 2013.

[18] Y. Mo, E. Garone, A. Casavola, and B. Sinopoli, “False data injection attacks against state estimation in wireless sensor networks,” in Decision and Control (CDC), 2010 49th IEEE Conference on. IEEE, 2010, pp. 5967–5972.

[19] C.-Z. Bai, F. Pasqualetti, and V. Gupta, “Data-injection attacks in stochastic control systems: Detectability and performance tradeoffs,” Automatica, vol. 82, pp. 251–260, 2017.

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

[21] H. Fawzi, P. Tabuada, and S. Diggavi, “Secure estimation and control for cyber-physical systems under adversarial attacks,” IEEE Transactions on Automatic Control, vol. 59, no. 6, pp. 1454–1467, 2014.

[22] Y. Shoukry, P. Martin, Y. Yona, S. Diggavi, and M. Srivastava, “Pyra: Physical challenge-response authentication for active sensors under spoofing attacks,” in Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security. ACM, 2015, pp. 1004–1015.

[23] S. Amin, A. A. Cárdenas, and S. S. Sastry, “Safe and secure networked control systems under denial-of-service attacks,” in International Workshop on Hybrid Systems: Computation and Control. Springer, 2009, pp. 31–45.

[24] A. Cetinkaya, H. Ishii, and T. Hayakawa, “Networked control under random and malicious packet losses,” IEEE Transactions on Automatic Control, vol. 62, no. 5, pp. 2434–2449, 2017.

[25] V. Dolk, P. Tesi, C. De Persis, and W. Heemels, “Event-triggered control synthesis under deception-of-service attacks,” IEEE Transactions on Control of Network Systems, vol. 4, no. 1, pp. 93–105, 2017.

[26] H. S. Foroush and S. Martinez, “On event-triggered control of linear stochastic systems under periodic denial-of-service jamming attacks,” in Decision and Control (CDC), 2012 IEEE 51st Annual Conference on. IEEE, 2012, pp. 2551–2556.

[27] Y. Shoukry, M. Chong, M. Wakaiki, P. Nuzzo, A. Sangiovanni-Vincentelli, A. S. Seshia, J. P. Hespanha, and P. Tabuada, “Smrt-based observer design for cyber-physical systems under sensor attacks,” ACM Transactions on Cyber-Physical Systems, vol. 2, no. 1, p. 5, 2018.

[28] N. Bezzo, J. Weimer, M. Pajic, O. Sokolsky, G. J. Pappas, and I. Lee, “Attack resilient state estimation for autonomous robotic systems,” in Intelligent Robots and Systems (IROS 2014), 2014 IEEE/RSJ International Conference on. IEEE, 2014, pp. 3692–3698.

[29] B. Satchidanandan and P. Kumar, “Control systems under attack,” in 2010 49th Annual Allerton Conference on. IEEE, 2010, pp. 107–114.

[30] T. Sarkar and A. Rakhlin, “How fast can linear dynamical systems be learned?” arXiv preprint arXiv:1812.01251, 2018.

[31] A. Rantzer, “Concentration bounds for single parameter adaptive control,” in 2018 Annual American Control Conference (ACC). IEEE, 2018, pp. 1862–1866.

[32] M. Simchowitz, H. Man, S. Tu, M. I. Jordan, and B. Recht, “Learning without mixing: Towards a sharp analysis of linear system identification,” arXiv preprint arXiv:1802.08334, 2018.

[33] G. Zames, “Adaptive control: Towards a complexity-based general theory,” Automatica, vol. 41, no. 2, pp. 1161–1167, 1998.

[34] K. J. Aström and B. Wittenmark, Adaptive control. Courier Corporation, 2013.

[35] R. S. Sutton and A. G. Barto, Reinforcement learning: An introduction. MIT press Cambridge, 1998, vol. 1, no. 1.

[36] P. Auer, N. Cesa-Bianchi, Y. Freund, and R. E. Schapire, “Gambling in a rigged casino: The adversarial multi-armed bandit problem,” in Foundations of Computer Science, 1995. Proceedings., 36th Annual Symposium on. IEEE, 1995, pp. 32–43.

[37] D. Libezon, Calculus of variations and optimal control theory: a concise introduction. Princeton University Press, 2011.

[38] E. L. Lehmann and J. P. Romano, Testing statistical hypotheses, 3rd ed. New York: NY: Springer Science & Business Media Inc., 2005.

[39] M. Amini and M. Fu, “Ergodic properties for multirate linear systems,” IEEE transactions on signal processing, vol. 55, no. 2, pp. 461–473, 2007.

[40] H. W. Sorensen, “Least-squares estimation: from Gauss to Kalman,” IEEE spectrum, vol. 7, no. 7, pp. 63–68, 1970.

[41] L. Ljung, “Consistency of the least-squares identification method,” IEEE Transactions on Automatic Control, vol. 21, no. 5, pp. 779–781, 1976.

[42] T. L. Lai and C. Z. Wei, “Least squares estimates in stochastic regression models with applications to identification and control of dynamic systems,” The Annals of Statistics, pp. 154–166, 1982.

[43] M. Raginsky, “Divergence-based characterization of fundamental limitations of adaptive dynamical systems,” in Communication, Control, and Computing (Allerton), 2010 48th Annual Allerton Conference on. IEEE, 2010, pp. 107–114.

[44] Y. Yang and A. Barron, “Information-theoretic determination of minimax rates of convergence,” Annals of Statistics, pp. 1564–1599, 1999.