Learning-Based Attacks in Cyber-Physical Systems

Mohammad Javad Khojasteh, *Student Member, IEEE*, Anatoly Khina, *Member, IEEE*, Massimo Franceschetti, *Fellow, IEEE*, and Tara Javidi, *Member, IEEE*

***Abstract*—We introduce the problem of learning-based attacks in a simple abstraction of cyber-physical systems—the case of a discrete-time, linear, time-invariant plant that may be subject to an attack that overrides sensor readings and controller actions. The attacker attempts to learn the dynamics of the plant and subsequently overrides the controller's actuation signal to destroy the plant without being detected. The attacker can feed fictitious sensor readings to the controller using its estimate of the plant dynamics and mimic the legitimate plant operation. The controller, in contrast, is constantly on the lookout for an attack; once the controller detects an attack, it immediately shuts the plant off. In the case of scalar plants, we derive an upper bound on the attacker's deception probability for any measurable control policy when the attacker uses an arbitrary learning algorithm to estimate the system dynamics. We then derive lower bounds for the attacker's deception probability for both scalar and vector plants by assuming an authentication test that inspects the empirical variance of the system disturbance. We also show how the controller can improve the security of the system by superimposing a carefully crafted privacy-enhancing signal on top of the "nominal control policy." Finally, for nonlinear scalar dynamics that belong to the reproducing kernel Hilbert space, we investigate the performance of attacks based on nonlinear Gaussian process learning algorithms.**

***Index Terms*—Cyber-physical system security, learning for dynamics and control, man-in-the-middle attack, physical layer authentication, secure control, system identification.**

Manuscript received June 27, 2020; revised September 6, 2020; accepted September 13, 2020. Date of publication September 30, 2020; date of current version February 26, 2021. This work was supported in part by the National Science Foundation under Award CNS-1446891 and Award ECCS-1917177, and in part by the European Union's Horizon 2020 Research and Innovation Program under the Marie Skłodowska-Curie Grant 708932. This article was presented in part at the 8th IFAC Workshop on Distributed Estimation and Control in Networked Systems, 2019 [1]. Recommended by Associate Editor S. Dey. *(Corresponding author: Mohammad Javad Khojasteh.)*

Mohammad Javad Khojasteh was with the Department of Electrical and Computer Engineering, University of California at San Diego, La Jolla, CA 92093 USA. He is now with the Center for Autonomous Systems and Technologies, California Institute of Technology, Pasadena, CA 91125 USA (e-mail: mjkhojas@caltech.edu).

Anatoly Khina is with the School of Electrical Engineering, Tel Aviv University, Tel Aviv 6997801, Israel (e-mail: anatolyk@tauex.tau.ac.il).

Massimo Franceschetti and Tara Javidi are with the Department of Electrical and Computer Engineering, University of California at San Diego, La Jolla, CA 92093 USA (e-mail: massimo@ece.ucsd.edu; tara@ece.ucsd.edu).

Digital Object Identifier 10.1109/TCNS.2020.3028035

I. Introduction

RECENT technological advances in wireless communications and computation, and their integration into networked control and cyber-physical systems (CPS), open the door to a myriad of new and exciting applications, including cloud robotics and automation [2]. However, the distributed nature of CPS is often a source of vulnerability. Security breaches in CPS can have catastrophic consequences ranging from hampering the economy by obtaining financial gain, to hijacking autonomous vehicles and drones, to terrorism by manipulating life-critical infrastructures [3]–[5]. Real-world instances of security breaches in CPS, that were discovered and made public, include the revenge sewage attack in Maroochy Shire, Australia; the Ukraine power grid cyber-attack; the German steel mill cyber-attack; the Davis-Besse nuclear power plant attack in Ohio, USA; and the Iranian uranium-enrichment facility attack via the Stuxnet malware [6]. Studying and preventing such security breaches via control-theoretic methods has received a great deal of attention in recent years [7]–[24].

An important and widely studied class of attacks in CPS is based on the "man-in-the-middle" (MITM) paradigm [25]: an attacker overrides the sensor signals transmitted from the physical plant to the controller with fake signals that mimic stable and safe operation. At the same time, the attacker also overrides the control signals with malicious inputs to push the plant toward a catastrophic trajectory. It follows that CPS must constantly monitor the plant outputs and look for anomalies in the fake sensor signals to detect such attacks. The attacker, in contrast, aims to generate fake sensor readings in a way that would be indistinguishable from the legitimate ones.

The MITM attack has been extensively studied in two special cases [25]–[29]. The first case is the *replay attack*, in which the attacker observes and records the legitimate system behavior for a given time window and then replays this recording periodically at the controller's input [26]–[28]. The second case is the *statistical-duplicate attack*, which assumes that the attacker has acquired complete knowledge of the dynamics and parameters of the system and can construct arbitrarily long fictitious sensor readings that are statistically identical to the actual signals [25], [29], [30]. The replay attack assumes no knowledge of the system parameters—and, as a consequence, it is relatively easy to detect. An effective way to counter the replay attack consists of superimposing a random watermark signal, unknown to the attacker, on top of the control signal

[30]–[34]. 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 on the attacker or controller behavior to ensure that it can be detected. To combat the attacker's full knowledge, the controller may adopt *moving target* [35]–[38] or *baiting* [39], [40] techniques. Alternatively, the controller may introduce private randomness in the control input using *watermarking* [29]. In this scenario, a vital assumption is made: although the attacker observes the true sensor readings, it is barred from observing the control actions, as otherwise, it would be omniscient and undetectable.

Our *contributions* are as follows. First, we observe that in many practical situations, the attacker does not have full knowledge of the system and cannot simulate a statistically indistinguishable copy of the system. In contrast, the attacker can carry out more sophisticated attacks than simply replaying previous sensor readings, by attempting to "learn" the system dynamics from the observations. For this reason, we study *learning-based attacks*, in which the attacker attempts to learn a model of the plant dynamics, and show that they can outperform replay attacks on linear systems by providing a lower bound on the attacker's deception probability using a simple learning algorithm. Second, we derive a converse bound on the attacker's deception probability in the special case of scalar systems. This holds for *any* (measurable) control policy and for *any* learning algorithm that may be used by the attacker to estimate the dynamics of the plant. These contributions regard the possibility of performing learning-based attacks. Another contribution regards the way to defend the system against these attacks. For any learning algorithm utilized by the attacker to estimate the dynamics of the plant, we show that adding a proper *privacy-enhancing signal* to the "nominal control policy" can lower the deception probability. Finally, we offer a treatment for nonlinear scalar dynamics that belong to a reproducing kernel Hilbert space, by studying the performance of a nonlinear attack based on machine-learning Gaussian process algorithms.

Throughout this article, we assume that the attacker has full access to both sensor and control signals. The controller, in contrast, has perfect knowledge of the system dynamics and tries to discover the attack from the observations that are maliciously injected by the attacker. This assumed information-pattern *imbalance* between the controller and the attacker is justified, since the controller is *tuned* much longer than the attacker and thus has knowledge of the system dynamics to a far greater precision than the attacker. In contrast, the attacker can completely hijack the sensor and control signals that travel through a communication network that has been compromised. Previous watermarking techniques [26], [29], [30] were only effective at securing the system if the attacker has no access to the control signals, which is not the case here. In contrast, since in our case, the attacker does not have full knowledge of the system dynamics, our privacy-enhancing signal is used to hamper the learning process of the attacker during the learning phase, rather than providing a unique signature to the control signal as in the case of watermarking.

Since in our setting the success or failure of the attacker is dictated by its learning capabilities, our work is also related to recent studies in learning-based control [41]–[50]. In contrast to these works, where tools developed in machine learning are used to design controllers in the presence of uncertainty, our work assumes a setting in which the controller has perfect knowledge of the system dynamics and tries to discover a possible attack from the observations. At the same time, the attacker aims to learn the system dynamics to construct a carefully crafted fictitious sensor reading signal to fool the controller. Thus, the security guarantees in this work are achieved by analyzing the performance and limitations of learning algorithms.

Learning-based attacks are also related to the known-plaintext attacks (KPA), introduced in [51], in linear systems with linear controllers. Using pole–zero analysis from classical system identification, Yuan and Mo [51] investigated necessary and sufficient conditions for which the system is identifiable by an attacker and, as a result, vulnerable against KPA. To combat KPA, the work [51] utilized low-rank linear controllers that trade control performance for security.

The outline of the rest of this article is as follows. The notations used in this work are detailed in Section I-A. For ease of exposition, we start by presenting the problem for the special case of scalar linear plants in Section II and present our main results for this case in Section III. We then extend the model and treatment to vector linear and scalar nonlinear plants in Section IV and Appendix A available online in [52], respectively. We conclude this article in Section V with a discussion of future directions. Due to space constraints, the appendixes (and some of the proofs) are available online in [52].

A. Notation

Throughout this article, we use the following notation. We denote by $\mathbb{N}$ the set of natural numbers and by $\mathbb{R}$ the set of real numbers. All logarithms, denoted by $\log$, are base 2. We denote by $\|\cdot\|$ the Euclidean norm of a vector and by $\|\cdot\|_{\text{op}}$ the *operator* norm induced by it when applied to a matrix. We denote by $\dagger$ the transpose operation of a matrix. For two real-valued functions g and h, $g(x) = O(h(x))$ as $x \to x_0$ means $\limsup_{x\to x_0} |g(x)/h(x)| < \infty$, and $g(x) = o(h(x))$ as $x \to x_0$ means $\lim_{x\to x_0} |g(x)/h(x)| = 0$. We denote by $x_i^j = (x_i, \ldots, x_j)$ the realization of the tuple of random variables $X_i^j = (X_i, \ldots, X_j)$ for $i, j \in \mathbb{N}, i \le j$. Random matrices are represented by boldface capital letters (e.g., $\mathbf{A}$), and their realizations are represented by typewriter boldface letters (e.g., $\mathtt{A}$). $\mathbf{A} \succeq \mathbf{B}$ means that $\mathbf{A} - \mathbf{B}$ is a positive-semidefinite matrix, namely $\succeq$ is the Löwner order of Hermitian matrices. $\lambda_{\max}(\mathbf{A})$ denotes the largest eigenvalue of the matrix $\mathbf{A}$. We represent the random vector with boldface small letters, and $\mathbf{x}_i^j = (\mathbf{x}_i, \ldots, \mathbf{x}_j)$ for $i, j \in \mathbb{N}, i \le j$. $\mathbb{P}_{\mathbf{x}}$ denotes the distribution of the random vector $\mathbf{x}$ with respect to (w.r.t.) probability measure $\mathbb{P}$, whereas $f_{\mathbf{x}}$ denotes its probability density function (PDF) w.r.t. to the Lebesgue measure, if it has one. An event is said to happen almost surely (a.s.) if it occurs with probability 1. For real numbers a and b, $a \ll b$ means a is much less than b, in some numerical

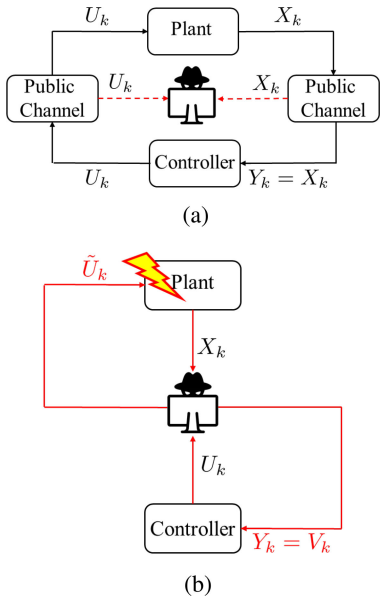

Fig. 1. System model during learning-based attack phases. (a) *Learning:* During this phase, the attacker eavesdrops and learns the system, without altering the input signal to the controller ($Y_k = X_k$). (b) *Hijacking:* During this phase, the attacker hijacks the system and intervenes as a MITM in two places: acting as a fake plant for the controller ($Y_k = V_k$) by impersonating the legitimate sensor, and as a malicious controller ($\tilde{U}_k$) for the plant aiming to destroy the plant.

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

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

where $E_{\mathbb{P}_X}$ denotes the expectation w.r.t. probability measure $\mathbb{P}_X$. The conditional KL divergence between probability distributions $\mathbb{P}_{Y|X}$ and $\mathbb{Q}_{Y|X}$ averaged over $\mathbb{P}_X$ is defined as $D(\mathbb{P}_{X|Y} \| \mathbb{Q}_{Y|X} | \mathbb{P}_X) \triangleq \mathbb{E}_{\mathbb{P}_{\tilde{X}}}[D(\mathbb{P}_{Y|X=\tilde{X}} \| \mathbb{Q}_{Y|X=\tilde{X}})]$, where $(X, \tilde{X})$ are independent and identically distributed (i.i.d.). The mutual information between random variables X and Y is defined as $I(X;Y) \triangleq D(\mathbb{P}_{XY} \| \mathbb{P}_X \mathbb{P}_Y)$. The conditional mutual information between random variables X and Y given random variable Z is defined as $I(X;Y|Z) \triangleq \mathbb{E}_{\mathbb{P}_{\tilde{Z}}}[I(X;Y|Z = \tilde{Z})]$, where $(Z, \tilde{Z})$ 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 system

$$X_{k+1} = aX_k + U_k + W_k \tag{1}$$

where X_k, a, U_k, and 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_1^k. If the attacker does not tamper sensor reading, at any time $k \in \mathbb{N}$, we have $Y_k = X_k$. We assume that the initial condition X_0 has a known (to all parties) distribution and is independent of the disturbance sequence $\{W_k\}$. For analytical purposes, we assume that the process $\{W_k\}$ has i.i.d. Gaussian samples of zero mean and variance σ^2 that is known to all parties. We assume, without loss of generality, that $W_0 = 0$ and $\mathbb{E}[X_0] = 0$, and take $U_0 = 0$. Moreover, to simplify the notation, let $Z_k \triangleq (X_k, U_k)$ denote the state-and-control input at time k and its trajectory up to time k by

$$Z_1^k \triangleq (X_1^k, U_1^k).$$

The controller is equipped with a detector that tests for anomalies in the observed history Y_1^k. When the controller detects an attack, it shuts the system down and prevents the attacker from causing further "damage" to the plant. The controller/detector is aware of the plant dynamics (1) and knows the open-loop gain a of the plant. In contrast, the attacker knows the plant dynamics (1) as well as the plant state X_k and control input U_k (or equivalently, Z_k) at time k (see Fig. 1). However, it does not know the open-loop gain a.

We assume that the open-loop gain is fixed in time, but unknown to the attacker (as in the frequentist approach [53]). Nevertheless, it will be convenient, for algorithm aid, to assume a prior over the open-loop gain of the plant and treat it as a random variable A, that is *fixed across time*, whose PDF f_A is known to the attacker and whose realization a is known to the controller (cf. Section V-D). We assume all random variables to exist on a common probability space with probability measure $\mathbb{P}$ and U_k to be a *measurable function* of Y_1^k for all time $k \in \mathbb{N}$. We also denote the probability measure conditioned on $A = a$ by $\mathbb{P}_a$. Namely, for any measurable event C, we define

$$\mathbb{P}_a(C) = \mathbb{P}(C|A = a).$$

A is assumed to be independent of X_0 and $\{W_k | k \in \mathbb{N}\}$.

A. Learning-Based Attacks

We now define *learning-based attacks* that consist of two disjoint, consecutive, passive, and active phases (cf. Section V-C).

Phase 1: Learning. During this phase, the attacker passively observes the control input and the plant state to learn the open-loop gain of the plant. As illustrated in Fig. 1(a), for all $k \in [0, L]$, the attacker observes the control input U_k and the plant state X_k and tries to learn the open-loop gain a, where L is the duration of the learning phase. We denote by $\hat{A}$ the attacker's estimate of the open-loop gain a. •

Phase 2: Hijacking. In this phase, the attacker aims to destroy the plant via $\tilde{U}_k$ while remaining undetected. As illustrated in Fig. 1(b), from time $L + 1$ and onward, the attacker hijacks the system and feeds a malicious control signal $\tilde{U}_k$ to the plant and a fictitious sensor reading $Y_k = V_k$ to the controller. •

We assume that the attacker can use *any arbitrary* learning algorithm to estimate the open-loop gain a during the learning phase, and when the estimation is completed, we assume that during the hijacking phase, the fictitious sensor reading is constructed, in a model-based manner (cf. Section V-B) as follows:

$$V_{k+1} = \hat{A}V_k + U_k + \tilde{W}_k\,, \qquad k = L, \ldots, T-1 \tag{2}$$

where $\tilde{W}_k$ for $k = L, \ldots, T-1$ are i.i.d. Gaussian $\mathcal{N}(0, \sigma^2)$; 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.

B. Detection

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 the legitimate system operation (corresponding to the *null hypothesis*), the controller observation Y_k behaves according to

$$Y_{k+1} - aY_k - U_k(Y_1^k) \sim \text{i.i.d. } \mathcal{N}(0, \sigma^2). \tag{3}$$

In the case of an attack, during Phase 2 ($k > L$), (3) can be rewritten as

$$V_{k+1} - aV_k - U_k = V_{k+1} - aV_k + \hat{A}V_k - \hat{A}V_k - U_k \tag{4a}$$

$$= \tilde{W}_k + \left(\hat{A} - a\right) V_k \tag{4b}$$

where (4b) follows from (2). Hence, the estimation error $(\hat{A} - a)$ dictates the ease with which an attack can be detected.

Since the Gaussian PDF with zero mean is fully characterized by its variance, we shall follow [29] and test for anomalies in the latter, i.e., test whether the empirical variance of (3) is equal to the second moment of the plant disturbance $\mathbb{E}[W^2]$. To that end, we shall use a test that sets a confidence interval of length $2\delta > 0$ around the expected variance, i.e., it checks whether

$$\frac{1}{T}\sum_{k=1}^{T}\left[Y_{k+1} - aY_k - U_k(Y_1^k)\right]^2 \in (\text{Var}[W] - \delta, \text{Var}[W] + \delta) \tag{5}$$

where T is called the *test time*. That is, as implied by (4), the attacker deceives the controller and remains undetected if

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

C. Performance Measures

Definition 1: The hijack indicator at test time T is defined as

$$\Theta_T \triangleq \begin{cases} 0 & \forall j \le T:\ Y_j = X_j \\ 1, & \text{otherwise.} \end{cases}$$

Θ_T is an oracle, and at the test time T the controller uses Y_1^T to construct an estimate $\hat{\Theta}_T$ of Θ_T. More precisely, $\hat{\Theta}_T = 0$ if (5) occurs; otherwise, $\hat{\Theta}_T = 1$. •

Definition 2: The probability of deception is the probability of the attacker deceiving the controller and remaining undetected at the time instant T

$$P_{\text{Dec}}^{a,T} \triangleq \mathbb{P}_a\left(\hat{\Theta}_T = 0 | \Theta_T = 1\right). \tag{6}$$

The detection probability at test time T is defined as

$$P_{\text{Det}}^{a,T} \triangleq 1 - P_{\text{Dec}}^{a,T}.$$

Likewise, the probability of false alarm is the probability of detecting the attacker when it is not present, namely

$$P_{\text{FA}}^{a,T} \triangleq \mathbb{P}_a\left(\hat{\Theta}_T = 1 | \Theta_T = 0\right). \quad \bullet$$

Applying Chebyshev's inequality to (5) and noting that the system disturbances are i.i.d. Gaussian of variance σ^2, we have

$$P_{\text{FA}}^{T} \le \frac{\text{Var}[W^2]}{\delta^2 T} = \frac{3\sigma^4}{\delta^2 T}.$$

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_{Dec}^{T}, P_{Det}^{T}, and P_{FA}^{T}, respectively. For instance, P_{Det}^{T} is defined, w.r.t. a PDF f_A of A, as

$$P_{\text{Det}}^{T} \triangleq \mathbb{P}\left(\hat{\Theta}_T = 1 | \Theta_T = 1\right) = \int_{-\infty}^{\infty} P_{\text{Det}}^{a,T}\, f_A(a) da. \tag{7}$$

III. Statement of the Results

In this section, we describe our main results for the case of scalar plants. We provide lower and upper bounds on the deception probability (6) of the learning-based attack (2), where the estimate $\hat{A}$ in (2) may be constructed using an *arbitrary* learning algorithm. Our results are valid for *any measurable* control policy U_k. We find a lower bound on the deception probability by characterizing what the attacker can at least achieve using a least-squares (LS) algorithm and derive an information-theoretic converse for any learning algorithm using Fano's inequality [54, Chs. 2.10 and 7.9]. While our analysis is restricted to the asymptotic case, $T \to \infty$, it is straightforward to extend it to the nonasymptotic case.

For analytical purposes, we assume that the power of the fictitious sensor reading is equal to $\beta^{-1} < \infty$, namely

$$\lim_{T\to\infty} \frac{1}{T}\sum_{k=L+1}^{T} V_k^2 = 1/\beta \qquad \text{a.s. w.r.t. } \mathbb{P}_a. \tag{8}$$

Remark 1: Assuming that the control policy is memoryless, namely, U_k is only dependent on Y_k, the process V_k is Markov for $k \ge L+1$. By further assuming that $L = o(T)$ and using the generalization of the law of large numbers for Markov processes [55], we deduce

$$\lim_{T\to\infty} \frac{1}{T}\sum_{k=L+1}^{T} V_k^2 \ge \text{Var}[W] \qquad \text{a.s. w.r.t. } \mathbb{P}_a.$$

Hence, in this case, we have $\beta \le 1/\text{Var}[W]$. Also, when the control policy is linear and stabilizes (2), that is, $U_k = -\Omega Y_k$ and $|\hat{A} - \Omega| < 1$, it is easy to verify that (8) holds true for $\beta = (1 - (\hat{A} - \Omega)^2)/\text{Var}[W]$. The assumption in (8) can also be relaxed as described in Remarks 4 and 6, in the following. •

In the following lemma, we show that for any learning-based attack (2), as $T \to \infty$, the empirical variance used in the variance test (5) can be expressed in terms of the estimation error. The result follows from the strong law of large numbers applied to

martingale difference sequences [56, Lemma 2, part iii]; it is proved in [52, App. C-A].

Lemma 1: Consider any learning-based attack (2) and any measurable control policy $\{U_k\}$ such that the fictitious sensor reading power satisfies (8). Then, the variance test (5) reduces a.s., w.r.t. $\mathbb{P}_a$, to

$$\lim_{T\to\infty}\frac{1}{T}\sum_{k=1}^{T}[Y_{k+1}-aY_k-U_k(Y_1^k)]^2=\mathrm{Var}\,[W]+\frac{(\hat{A}-a)^2}{\beta}.$$

A. Lower Bound on the Deception Probability

To provide a lower bound on the deception probability $P_{\mathrm{Dec}}^{a,T}$, we consider a specific estimate $\hat{A}$ at the conclusion of the first phase by the attacker. Namely, we use LS estimation due to its efficiency and amenability to recursive update over observed incremental data [44]–[46]. The LS algorithm approximates the overdetermined system of equations

$$\begin{pmatrix}X_2\\X_3\\\vdots\\X_L\end{pmatrix}=A\begin{pmatrix}X_1\\X_2\\\vdots\\X_{L-1}\end{pmatrix}+\begin{pmatrix}U_1\\U_2\\\vdots\\U_{L-1}\end{pmatrix}$$

by minimizing the Euclidean distance $\hat{A}=\mathrm{argmin}_A\,\|X_{k+1}-AX_k-U_k\|$ to estimate (or "identify") the plant, the solution to which is

$$\hat{A}=\frac{\sum_{k=1}^{L-1}(X_{k+1}-U_k)X_k}{\sum_{k=1}^{L-1}X_k^2}\qquad\text{a.s. w.r.t. }\mathbb{P}_a.\tag{9}$$

Remark 2: Equation (9) is well defined since $\mathbb{P}_a(X_k=0)=0$, as we assumed that W_k are i.i.d. zero-mean Gaussian for all $k\in\mathbb{N}$. •

We now lower bound the deception probability of an attacker that utilizes LS estimation (9) under the variance test and in the presence of any measurable control policy, for which (8) holds. The following theorem demonstrates the *existence* of a learning-based attack that satisfies this lower bound. As other learning algorithms may lead to better estimates, this also serves as a lower bound on the attacker's deception probability in the general case.

Theorem 1: Consider LS (9) learning-based attack (2) and any measurable control policy $\{U_k\}$ such that the fictitious sensor readings satisfy (8). Then, the asymptotic deception probability under the variance test (5) is lower bounded as

$$\lim_{T\to\infty}P_{\mathrm{Dec}}^{a,T}=\mathbb{P}_a\left(|\hat{A}-a|<\sqrt{\delta\beta}\right)\tag{10a}$$

$$\geq\mathbb{P}_a\left(\frac{\left|\sum_{k=1}^{L-1}W_kX_k\right|}{\sum_{k=1}^{L-1}X_k^2}<\sqrt{\delta\beta}\right)\tag{10b}$$

$$\geq 1-\frac{2}{(1+\delta\beta)^{L/2}}.\tag{10c}$$

Proof: Equation (10a) follows from Lemma 1 and the dominated convergence theorem [55]. For details, see [52, Appendix C-B]. Clearly, the estimation error of the LS algorithm (9) is [44]

$$\hat{A}-a=\frac{\sum_{k=1}^{L-1}W_kX_k}{\sum_{k=1}^{L-1}X_k^2}\qquad\text{a.s. w.r.t. }\mathbb{P}_a.\tag{11}$$

Consequently, by (10a), a learning-based attack (2) can at least achieve the asymptotic deception probability (10b). Finally, (10c) holds by the concentration of measure [44, Theor. 4] by noting that U_k is a measurable function of $Y_1^k=X_1^k$, for $k\in\{1,\ldots,L\}$.

Remark 3: We can study the special case of Theorem 1 for a linear controller. Using the value of β calculated in Remark 1 for a linear controller $U_k=-\Omega Y_k$ when $|\hat{A}-\Omega|<1$, we can rewrite (10c) as

$$\lim_{T\to\infty}P_{\mathrm{Dec}}^{a,T}\geq 1-\frac{2}{\left(1+\delta\frac{1-(\hat{A}-\Omega)^2}{\mathrm{Var}[W]}\right)^{L/2}}.\tag{12}$$

In this case, for a fixed L, as $\mathrm{Var}[W]$ increases, the lower bound in (12) decreases. The reduction of the attacker's success rate can be explained by noticing that LS estimation (9) is based on minimizing $\|X_{k+1}-AX_k-U_k\|$, and the precision of this estimate decreases as $\mathrm{Var}[W]$ increases. •

Remark 4: By replacing the limit with limsup in the lower bound in Theorem 1, the result holds even if the limit in (8) does not exist. Also, if the limit in (8) is infinite, then either the attacker or the controller is doing a poor job, as described next. Assume that the attacker uses an estimate $\hat{A}$ such that at the conclusion of the learning phase $\hat{A}$ belongs to the interval $(A-\delta',A+\delta')$, for a small value of $\delta'>0$. In this case, if the power of the fictitious sensor reading tends to infinity, then the controller is not robust. In contrast, assume that the controller stabilizes the system with any open-loop gain that belongs to the interval $(A-\delta',A+\delta')$ where $\delta'>0$. In this case, if the power of fictitious sensor tends to infinity then the attacker's estimate at the conclusion of the learning phase does not belong to the interval $(A-\delta',A+\delta')$, i.e., the absolute value of the attacker's estimation error is larger than δ'. •

Example 1: In this example, we compare the empirical performance of the variance test with our developed bound in Theorem 1. At every time T, the controller tests the empirical variance for anomalies over a detection window $[1,T]$, using a confidence interval $2\delta>0$ around the expected variance (5). Here, $a=1,\delta=0.1,U_k=-0.88aY_k$ for all $1\leq k\leq T=800$ and $\{W_k\}$ are i.i.d. Gaussian $\mathcal{N}(0,1)$, and 500 Monte Carlo simulations were performed.

The learning-based attacker (2) uses the LS algorithm (9) to estimate a and as illustrated in Fig. 2, the attacker's success rate increases as the duration of the learning phase L increases. This is in agreement with (10c) since the attacker can improve its estimate of a and the estimation error $|\hat{A}-a|$ reduces as L increases. As discussed in Section II-C, 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, the attacker's success rate could potentially tend to 1. Indeed, such behavior is observed in Fig. 2 for a learning-based attacker (2)

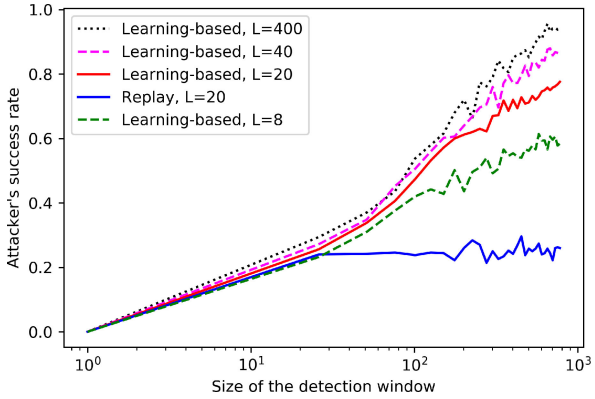

Fig. 2. Attacker's success rate $P_{\text{Dec}}^{a,T}$ versus the size of the detection window T.

with $L = 400$. Fig. 2 also illustrates that our learning-based attack outperforms the replay attack. A replay attack with a recording length of $L = 20$ and a learning-based attack with a learning phase of length $L = 20$ are compared, and the success rate of the replay attack saturates at a lower value. Moreover, a learning-based attack with a learning phase of length $L = 8$ has a higher success rate than a replay attack with a larger recording length of $L = 20$. •

B. Upper Bound on the Deception Probability

We now derive an upper bound on the deception probability (6) of any learning-based attack (2) where $\hat{A}$ in (2) is constructed using *any arbitrary* learning algorithm, for *any measurable* control policy, when A is distributed over a symmetric interval $[-R, R]$. Since the uniform distribution has the highest entropy among all distributions with finite support [54, Ch. 12], we assume that A has a uniform prior over the interval $[-R, R]$. We further assume that the attacker knows this distribution (including the value of R), whereas the controller knows the true realization of A (as before).

Theorem 2: For any $R > 0$, let A be distributed uniformly over $[-R, R]$ and consider *any measurable* control policy $\{U_k\}$, and *any learning-based attack* (2) such that the fictitious sensor readings satisfy (8) with $\sqrt{\delta\beta} \leq R$. Then, the asymptotic deception probability when using the variance test (5), is upper bounded as

$$\lim_{T\to\infty} P_{\text{Dec}}^{T} = \mathbb{P}(|A - \hat{A}| < \sqrt{\delta\beta}) \tag{13a}$$

$$\leq \Lambda \triangleq \frac{I(A; Z_1^L) + 1}{\log(R/\sqrt{\delta\beta})}. \tag{13b}$$

In addition, if for all $k \in \{1, \dots, L\}$, $A \to (X_k, Z_1^{k-1}) \to U_k$ is a Markov chain, then for any sequence of probability measures $\{\mathbb{Q}_{X_k|Z_1^{k-1}}\}$, such that for all $k \in \{1, \dots, L\}$ $\mathbb{P}_{X_k|Z_1^{k-1}} \ll \mathbb{Q}_{X_k|Z_1^{k-1}}$, we have

$$\Lambda \leq \frac{\sum_{k=1}^{L} D\left(\mathbb{P}_{X_k|Z_1^{k-1},A} \| \mathbb{Q}_{X_k|Z_1^{k-1}} | \mathbb{P}_{Z_1^{k-1},A}\right) + 1}{\log\left(R/\sqrt{\delta\beta}\right)}. \tag{14}$$

Proof: Equation (13a) follows from (7), (10a), and Tonelli's theorem [55]. For details, see [52, Appendix D]. Equation (13b) follows by noting that since the attacker observes the plant state and control input during the learning phase, which lasts L steps, and since $A \to (X_1^L, U_1^L) \to \hat{A}$ constitutes a Markov chain, using the continuous domain version of Fano's inequality [57, Prop. 2], we have

$$\inf_{\hat{A}} \mathbb{P}\left(\left|\hat{A} - A\right| \geq \sqrt{\delta\beta}\right) \geq 1 - \frac{I(A; Z_1^L) + 1}{\log(R/\sqrt{\delta\beta})}$$

whenever $\sqrt{\delta\beta} \leq R$. Finally, (14) follows the arguments of [58] and is proven, for completeness, in [52, Appendix D-B]. ■

Remark 5: By looking at the numerator in (13b), it follows that the bound on the deception probability becomes looser as the amount of information revealed about the open-loop gain A by the observation Z_1^L increases. In contrast, by looking at the denominator, the bound becomes tighter as R increases. This is consistent with the observation of Zames [58] that system identification 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 poorer estimation of A by the attacker, which leads, in turn, to a decrease in the achievable deception probability. The denominator can also be interpreted as the intrinsic uncertainty of A when it is observed at resolution $\sqrt{\delta\beta}$ as it corresponds to the entropy of the random variable A when it is quantized at such resolution. •

In conclusion, Theorem 2 provides two upper bounds on the deception probability. The first bound (13b) 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_1^L during the exploration phase—reduces the deception probability. The second bound (14) allows freedom in choosing the auxiliary probability measure $\mathbb{Q}_{X_k|Z_1^{k-1}}$, making it a rather useful bound. For instance, by choosing $\mathbb{Q}_{X_k|Z_1^{k-1}} \sim \mathcal{N}(0, \sigma^2)$, for all $k \in \mathbb{N}$, we can rewrite the upper bound (14) in terms of $\mathbb{E}_{\mathbb{P}}[(AX_{k-1} + U_{k-1})^2]$ as follows. The proof of the following corollary can be found in [52].

Corollary 1: Under the assumptions of Theorem 2, if for all $k \in \{1, \dots, L\}$, $A \to (X_k, Z_1^{k-1}) \to U_k$ is a Markov chain, then the asymptotic deception probability upper bounded as

$$\lim_{T\to\infty} P_{\text{Dec}}^{T} \leq G(Z_1^L), \tag{15a}$$

$$G(Z_1^L) \triangleq \frac{\frac{\log e}{2\sigma^2} \sum_{k=1}^{L} \mathbb{E}_{\mathbb{P}}\left[(AX_{k-1} + U_{k-1})^2\right] + 1}{\log\left(R/\sqrt{\delta\beta}\right)}. \tag{15b}$$

Remark 6: Following the same discussion as in Remark 4, by replacing the limit with liminf in the upper bound in Theorem 2, the derived results remain true even if (8) does not happen. •

The next example compares the lower and upper bounds on the deception probability of Theorem 1 and Corollary 1.

Example 2: Theorem 1 provides a lower bound on the deception probability given $A = a$. Hence, by applying the law of total probability w.r.t. the PDF f_A as in (7), we can apply the result of Theorem 1 to provide a lower bound also on the average deception probability for a random open-loop gain A. In this

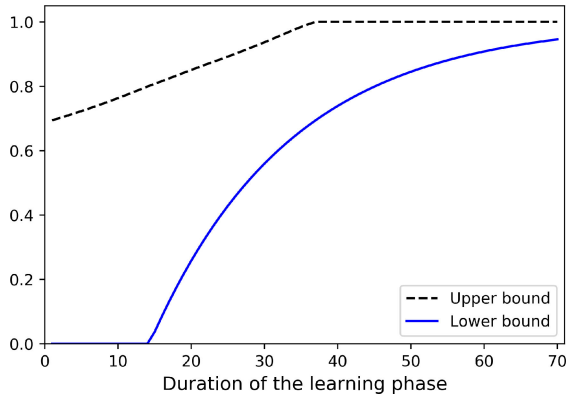

Fig. 3. Comparison of the lower and upper bounds on the deception probability of Theorem 1 and Corollary 1, respectively.

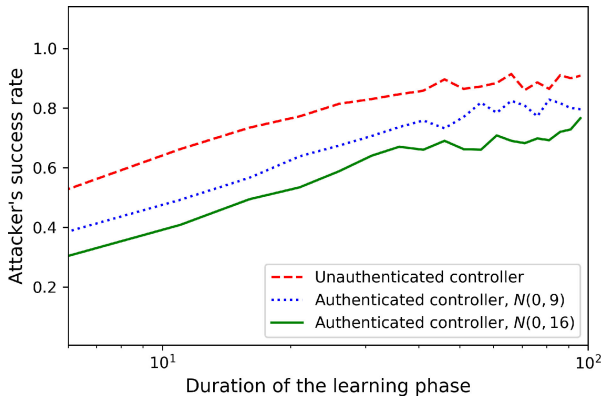

Fig. 4. Attacker's success rate $P_{\text{Dec}}^{a,T}$ versus the duration of the learning phase L.

context, Fig. 3 compares the lower and upper bounds on the deception probability provided by Theorem 1 and Corollary 1, augmented with the trivial cases of zero and one probability, namely, $\max\{0, 1-(2/(1+\delta\beta)^{L/2})\}$ and $\min\{1, G(Z_1^L)\}$ where A is distributed uniformly over $[-0.9, 0.9]$. Equation (16a) is valid when the control input is not a function of random variable A; hence, we assumed $U_k = -0.045Y_k$ for all time $k \in \mathbb{N}$. Here, $\delta = 0.1$, $\{W_k\}$ are i.i.d. Gaussian with zero mean and variance of 0.16 and for simplicity, we assume the limit in (8) exists [cf. Remarks 4 and 6] and we let $\beta = 1.1$. Although, in general, the attacker's estimation of the random open-loop gain A and consequently the power of fictitious sensor reading (8) vary based on the learning algorithm and the realization of A, the comparison of the lower and upper bounds in Fig. 3 is restricted to a fixed β. 2000 Monte Carlo simulations were performed.

Fig. 3 also illustrates the gap between these lower and upper bounds on the deception probability. By restricting the class of control policies or learning algorithms, one might be able to derive tighter results at the cost of losing generality. •

C. Privacy-Enhancing Signal

For a given duration of the learning phase L, to increase the security of the system, at any time k, the controller can add a privacy-enhancing signal Γ_k to an unauthenticated control policy $\{\bar{U}_k | k \in \mathbb{N}\}$:

$$U_k = \bar{U}_k + \Gamma_k\,, \qquad k \in \mathbb{N}. \tag{16}$$

We refer to such a control policy U_k as the *authenticated* control policy $\bar{U}_k$. We denote the states of the system that would be generated if only the unauthenticated control signal $\bar{U}_1^k$ were applied by $\bar{X}_1^k$ and the resulting trajectory by $\bar{Z}_1^k \triangleq (\bar{X}_1^k, \bar{U}_1^k)$.

The following numerical example illustrates the effect of the privacy-enhancing signal on the deception probability.

Example 3: Here, the attacker uses the LS algorithm (9), the detector uses the variance test (5), $a = 1, T = 600, \delta = 0.1$, and $\{W_k\}$ are i.i.d. standard Gaussian. We compare the attacker's success rate, the empirical $P_{\text{Dec}}^{a,T}$, as a function of the duration L of the learning phase for three different control policies: 1) unauthenticated control signal $\bar{U}_1^k = -aY_k$ for all k; 2) authenticated control signal (17), where Γ_k are i.i.d. Gaussian $\mathcal{N}(0, 9)$; and 3) authenticated control signal (16), where Γ_k are i.i.d. Gaussian $\mathcal{N}(0, 16)$. As illustrated in Fig. 4 , for both the authenticated and unauthenticated control signals, the attacker's success rate increases as the duration of the learning phase increases. This is in agreement with (10c) since the attacker can improve its estimate of a as L increases. Furthermore, for a fixed L, the attacker performance deteriorates as the power of privacy-enhancing signal Γ_k increases. Namely, Γ_k hampers the learning process of the attacker and the estimation error $|\hat{A} - a|$ increases as the power of the privacy-enhancing signal increases. 500 Monte Carlo simulations were performed. •

Remark 7: A "good" privacy-enhancing signal entails little increase in the control cost [59] compared to its unauthenticated version while providing enhanced detection probability (6) and/or false alarm probability. Finding the optimal privacy-enhancing signal is left for future research. We remark that since the submission of this article [1], [52], some latter literature has appeared that builds on it. In particular, in a follow-up work [60], Ziemann and Sandberg focused on designing optimal privacy-enhancing signal, by studying the optimal control problem of linear systems regularized with Fisher information, where the latter serves as a proxy to the estimation quality of A via the Cramer–Rao lower bound. •

One may envisage that superimposing any noisy signal Γ_k on top of the control policy $\{\bar{U}_k | k \in \mathbb{N}\}$ would necessarily enhance the detectability of *any* learning-based attack (2) since the observations of the attacker are in this case noisier. However, it turns out that injecting a strong noise for some learning algorithm may, in fact, speed up the learning process as it improves the power of the signal magnified by the open-loop gains w.r.t. the observed noise [44]. Any signal Γ_k that satisfies the condition proposed in the following corollary, whose proof available in [52], will provide enhanced guarantees on the detection probability when the attacker uses *any arbitrary* learning algorithm to estimate the uniformly distributed A over the symmetric interval $[-R, R]$.

Corollary 2: For any control policy $\{\bar{U}_k | k \in \mathbb{N}\}$ with trajectory $\bar{Z}_1^k = (\bar{X}_1^k, \bar{U}_1^k)$ and its corresponding authenticated control policy U_1^k (17) with trajectory $Z_1^k = (X_1^k, U_1^k)$, under

the assumptions of Corollary 1, if for all $k \in \{1, \dots, L-1\}$

$$\mathbb{E}_{\mathbb{P}}\left[\Psi_k^2 + 2\Psi_k(A\bar{X}_k + \bar{U}_k)\right] < 0 \tag{17}$$

where $\Psi_k \triangleq \sum_{j=1}^{k} A^{k-j}\Gamma_j$, for any $L \geq 2$, the following majorization of G (15b) holds:

$$G(Z_1^L) < G(\bar{Z}_1^L). \tag{18}$$

Remark 8: Corollary 2 can be generalized by replacing the limit with liminf in (8) (cf. Remarks 4 and 6). •

Example 4: In this example, we describe a class of privacy-enhancing signals that yield better guarantees on the deception probability. For all $k \in \{2, \dots, L\}$, clearly, $\Psi_{k-1} = -(A\bar{X}_{k-1} + \bar{U}_{k-1})/\eta$ satisfies the condition in (17) for any $\eta \geq 3$. Thus, by choosing the privacy-enhancing signals $\Gamma_1 = -(A\bar{X}_1 + \bar{U}_1)/\eta$, and $\Gamma_k = -(A\bar{X}_k + \bar{U}_k)/\eta - \sum_{j=1}^{k-1} A^{k-1-j}\Gamma_j$ for all $k \in \{3, \dots, L\}$, (18) holds. A numerical example for this authentication policy demonstrates a decrease in the deception probability at the expense of a higher control cost, and it can be found in [52, Appendix B-A]. •

Remark 9: The privacy-enhancing signal introduced in this work is related to the dynamic watermarking signal [26], [29], [30], which are unique signatures that are available *only* to the controller. In contrast, in our setup [depicted in Fig. 1(b)], the attacker has access to the signal generated by the controller. Thus, by reading the control input at time k and constructing the fictitious sensor reading V_{k+1}, as in (2), the attacker can construct a fictitious sensor reading containing any watermark signal inscribed by the controller. It follows that techniques based on dynamic watermarking that rely on the privacy of such signal break down in the case where attacks have access to the control signal generated by the controller. Instead, we take advantage of the authentication signal in a different way: since the attacker does not have full knowledge about the system dynamics, this signal is used to hamper the learning process of the attacker during the learning phase. •

IV. Extension to Vector Systems

We now generalize our results to vector systems. Consider the networked control system depicted in Fig. 1, with the plant dynamics replaced by a *vector* plant:

$$\mathbf{x}_{k+1} = \mathtt{A}\mathbf{x}_k + \mathbf{u}_k + \mathbf{w}_k \tag{19}$$

where $\mathbf{x}_k \in \mathbb{R}^{n\times 1}$, $\mathbf{u}_k \in \mathbb{R}^{n\times 1}$, $\mathtt{A} \in \mathbb{R}^{n\times n}$, and $\mathbf{w}_k \in \mathbb{R}^{n\times 1}$ represent the plant state, control input, open-loop gain of the plant, and plant disturbance, respectively, at time $k \in \mathbb{N}$. The controller, at time k, observes $\mathbf{y}_k$ and generates a control signal $\mathbf{u}_k$ as a function of $\mathbf{y}_1^k$, and $\mathbf{y}_k = \mathbf{x}_k$ at times $k \in \mathbb{N}$ at which the attacker does not tamper the sensor reading. We assume that the initial condition $\mathbf{x}_0$ has a known (to all parties) distribution and is independent of the disturbance sequence $\{\mathbf{w}_k\}$. For analytical purposes, we further assume that $\{\mathbf{w}_k\}$ is a process with i.i.d. multivariate Gaussian samples of zero mean and a covariance matrix $\mathbf{\Sigma}$ that is known to all parties. Without loss of generality, we assume that $\mathbf{w}_0 = 0$, $\mathbb{E}[\mathbf{x}_0] = 0$, and take $\mathbf{u}_0 = 0$.

We assume that the attacker uses the vector analog of learning-based attacks described in Section II-A, where the attacker can use *any* learning algorithm to estimate the open-loop gain matrix $\mathtt{A}$ during the learning phase. The estimation $\hat{\mathbf{A}}$ constructed by the attacker at the conclusion of the learning phase is utilized to construct the fictitious sensor readings $\{\mathbf{v}_k\}$ according to the vector analog of (2), where $\{\tilde{\mathbf{w}}_k | k = L, \dots, T-1\}$ are i.i.d. multivariate Gaussian with zero mean and covariance matrix $\mathbf{\Sigma}$.

Similar to the scalar case, for analytical purposes, we assume that the power of the fictitious sensor reading is equal to $1/\beta < \infty$ [cf. Remarks 1 and 4], namely

$$\lim_{T\to\infty} \frac{1}{T} \sum_{k=L+1}^{T} \|\mathbf{v}_k\|^2 = \frac{1}{\beta} \qquad \text{a.s. w.r.t. } \mathbb{P}_{\mathtt{A}}. \tag{20}$$

Since the zero-mean multivariate Gaussian distribution is completely characterized by its covariance matrix, we shall follow [29] and test for anomalies in the latter. To that end, define the error matrix

$$\mathbf{\Delta} \triangleq \mathbf{\Sigma} - \frac{1}{T} \sum_{k=1}^{T} [\mathbf{y}_{k+1} - \mathtt{A}\mathbf{y}_k - \mathbf{u}_k] [\mathbf{y}_{k+1} - \mathtt{A}\mathbf{y}_k - \mathbf{u}_k]^{\dagger}.$$

As in (5), we use a test that sets a confidence interval, w.r.t. the norm, around the expected covariance matrix, i.e., it checks whether

$$\|\mathbf{\Delta}\|_{\text{op}} \leq \gamma \tag{21}$$

at the test time T. For the sake of analysis, we use the operator norm in (21), which satisfies the submultiplicativity property.

The following lemma provides a necessary and sufficient condition for any learning-based attack [the vector analog of (2)] to deceive the controller and remain undetected, for a multivariate plant (19) under a covariance test (21), in the limit of $T \to \infty$; its proof is available in [52].

Lemma 2: Consider the multivariate plant (19) and any learning-based attack analogous to (2), with fictitious sensor reading power that satisfies (20), and any measurable control policy $\{\mathbf{u}_k\}$. Then, the attacker can deceive the controller and remain undetected, under the covariance test (21), a.s. in the limit $T \to \infty$, if and only if

$$\lim_{T\to\infty} \frac{1}{T} \left\| \sum_{k=L+1}^{T} (\hat{\mathbf{A}} - \mathtt{A})\mathbf{v}_k \mathbf{v}_k^{\dagger} (\hat{\mathbf{A}} - \mathtt{A})^{\dagger} \right\|_{\text{op}} \leq \gamma. \tag{22}$$

Lemma 2 has the following important implication:

$$\lim_{T\to\infty} \frac{1}{T} \left\| \sum_{k=L+1}^{T} (\hat{\mathbf{A}} - \mathtt{A})\mathbf{v}_k \mathbf{v}_k^{\dagger} (\hat{\mathbf{A}} - \mathtt{A})^{\dagger} \right\|_{\text{op}} \tag{23a}$$

$$\leq \lim_{T\to\infty} \frac{1}{T} \sum_{k=L+1}^{T} \left\| (\hat{\mathbf{A}} - \mathtt{A})\mathbf{v}_k ((\hat{\mathbf{A}} - \mathtt{A})\mathbf{v}_k)^{\dagger} \right\|_{\text{op}} \tag{23b}$$

$$\leq \lim_{T\to\infty} \frac{1}{T} \sum_{k=L+1}^{T} \left\| (\hat{\mathbf{A}} - \mathtt{A})\mathbf{v}_k \right\|_{\text{op}}^2 \tag{23c}$$

$$\leq \left\| \hat{\mathbf{A}} - \mathtt{A} \right\|_{\text{op}}^2 / \beta \tag{23d}$$

where (23b) follows from the triangle inequality, and (23c) and (23d) follow from the submultiplicativity of the operator, the

identity $\|\mathbf{v}_k\| = \|\mathbf{v}_k\|_{\text{op}}$, and by putting the power constraint (20) into force.

If $\|\hat{\mathbf{A}} - \mathtt{A}\|_{\text{op}}^2 \leq \gamma\beta$, (22) holds in the limit of $T \to \infty$, then the attacker is able to deceive the controller and remain undetected a.s., by Lemma 2. Equation (23) then implies that the norm of the estimation error, $\|\hat{\mathbf{A}} - \mathtt{A}\|_{\text{op}}$, dictates the ease with which an attack can go undetected. This is used next to develop a lower bound on the deception probability.

A. Lower Bound on the Deception Probability

We start by observing that in the case of multivariate systems, and in contrast to their scalar counterparts, some control actions might not reveal the entire plant dynamics, and in this case, the attacker might not be able to learn the plant completely. This phenomenon is captured by the persistent excitation property of control inputs, which describes control action signals that are sufficiently rich to excite all the system modes that will allow to learn them. While *avoiding* persistently exciting control inputs can be used as a way to *secure* the system against learning-based attacks, here, we assume a probabilistic variant of this property [58], [61].

Definition 3 (Persistent excitation): Given a plant (19), $\zeta > 0$, and $\rho \in [0, 1]$, the control policy $\mathbf{u}_k$ is (ζ, ρ)-persistently exciting if there exists a time $L_0 \in \mathbb{N}$ such that, for all $\tau \geq L_0$,

$$\mathbb{P}_{\mathtt{A}}\left(\frac{1}{\tau}\mathbf{G}_\tau \succeq \zeta \mathtt{I}_{n\times n}\right) \geq \rho, \tag{24}$$

where $\mathbf{G}_\tau$ is the sum of the state Gramians up to time τ:

$$\mathbf{G}_\tau \triangleq \sum_{k=1}^{\tau} \mathbf{x}_k \mathbf{x}_k^\dagger. \tag{25}$$

As in Section III-A, to find a lower bound on the deception probability $P_{\text{Dec}}^{\mathtt{A},T}$, we consider a specific estimate of $\mathtt{A}$, obtained via the LS estimation algorithm, analogous to (9), at the conclusion of the first phase by the attacker:

$$\hat{\mathbf{A}} = \begin{cases} \mathbf{0}_{n\times n}, & \det(\mathbf{G}_{L-1}) = 0 \\ \sum_{k=1}^{L-1}(\mathbf{x}_{k+1} - \mathbf{u}_k)\mathbf{x}_k^\dagger \mathbf{G}_{L-1}^{-1}, & \text{otherwise} \end{cases} \tag{26}$$

where $\mathbf{0}_{k\times\ell}$ denotes an all zero matrix of dimensions $k \times \ell$.

Next, we show an upper bound on the estimation error norm, $\|\hat{\mathbf{A}} - \mathtt{A}\|_{\text{op}}$, of the above LS algorithm (26) and use it to extend the bound in (10) to the vector case. A complete proof is available in [52].

Lemma 3: Consider the vector plant (19). If the attacker constructs $\hat{\mathbf{A}}$ using LS estimation (26) and the controller uses a policy $\{\mathbf{u}_k\}$ for which the event in (24) occurs for $\tau = L - 1$, that is, $\mathbf{G}_{L-1}/(L-1) \succeq \zeta \mathtt{I}_{n\times n}$, then we have

$$\|\hat{\mathbf{A}} - \mathtt{A}\|_{\text{op}} \leq \frac{1}{\zeta L}\sum_{k=1}^{L-1} \|\mathbf{w}_k \mathbf{x}_k^\dagger\|_{\text{op}} \qquad \text{a.s. w.r.t. } \mathbb{P}_{\mathtt{A}}. \tag{27}$$

The following theorem provides a lower bound on the deception probability of an attacker that utilizes LS estimation (26), and its proof can be found in [52]. As discussed before Theorem 1, since the attacker might be able to construct better estimates using other learning algorithms, this also serves as a lower bound on the attacker's deception probability in the general case.

Theorem 3: Consider the plant (19) with a (ζ, ρ)-persistently exciting control policy $\{\mathbf{U}_k\}$ from time L_0, and LS (26) learning-based attack [the vector analog of (2)] such that the fictitious sensor reading power satisfies (20) and with a learning phase of duration $L \geq L_0 + 1$. Then, the asymptotic deception probability, when using the covariance test (21), is bounded from below as

$$\lim_{T\to\infty} P_{\text{Dec}}^{\mathtt{A},T} \geq \mathbb{P}_{\mathtt{A}}\left(\|\hat{\mathbf{A}} - \mathtt{A}\|_{\text{op}} < \sqrt{\gamma\beta}\right) \tag{28a}$$

$$\geq \rho \mathbb{P}_{\mathtt{A}}\left(\frac{1}{\zeta L}\sum_{k=1}^{L-1} \|\mathbf{w}_k \mathbf{x}_k^\dagger\|_{\text{op}} < \sqrt{\gamma\beta}\right). \tag{28b}$$

Remark 10: The bound (10c) for scalar systems, which is *independent* of the control policy and state value, has been developed using the concentration bounds of [44] for the scalar LS algorithm (9). To the best of our knowledge, there are no similar concentration bounds for the vector variant of the LS algorithm (9) which work for *any* $\mathtt{A}$, and a large class of control policies. Looking for such bounds, which are independent of the state value, seems an interesting research venue. The lower bound (28b) is similar to (10b), while (10b) is stronger for the particular case of scalar system, as the upper bound on the estimation error derived in Lemma 3 is not required for the scalar case, and the estimation error is given in (11). •

Example 5: In this example, we compare the empirical performance of the covariance test against the learning-based attack, which utilizes LS estimation (26), and the replay attack. At every time k, the controller tests the empirical covariance for anomalies over a detection window $[1, T]$, using a confidence interval $2\gamma > 0$ around the operator norm of error matrix $\mathbf{\Delta}$ (21). Since we are considering the Euclidean norm for vectors, the induced operator norm amounts to $\|\mathbf{\Delta}\|_{\text{op}} = \sqrt{\lambda_{\max}(\mathbf{\Delta}^\dagger\mathbf{\Delta})}$. Here, $\gamma = 0.1$, $\mathbf{u}_k = -0.9\mathtt{A}\mathbf{y}_k$ for all $1 \leq k \leq T = 600$, and

$$\mathtt{A} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \qquad \mathbf{\Sigma} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}. \tag{29}$$

Fig. 5 presents the performance averaged over 180 runs of a Monte Carlo simulation. It illustrates that the vector variant of our learning-based attack also *outperforms* the replay attack. A learning-based attack with a learning phase of length $L = 40$ has a higher success rate than a replay attack with a larger recording length of $L = 50$. Similarly to the discussion for scalar systems in Section II-C, the false alarm rate decays to zero as the size of the detection window T tends to infinity. Thus, the success rate of learning-based attacks increases as the size of the detection window increases. Finally, as illustrated in Fig. 5, the attacker's success rate increases as the duration of the learning phase L increases, since the attacker improves its estimate of $\mathtt{A}$ as L increases. •

An example that investigates the effect of privacy-enhancing signals on the empirical deception probability of the learning-based attacks on the vector systems is in [52, Appendix B-B].

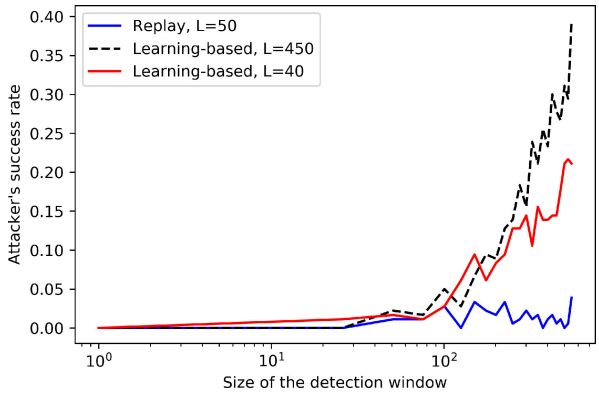

Fig. 5. Attacker's success rate $P_{\text{Dec}}^{\mathbb{A},T}$ versus the size of the detection window T.

V. Discussion and Future Work

A. Upper Bound on the Deception Probability

The upper bound in (23), which relates the deception criterion (22) to the estimation error $\|\hat{\mathbf{A}} - \mathbb{A}\|_{\text{op}}$, is used to find the lower bound (28). Finding a corresponding lower bound in term of $\|\hat{\mathbf{A}} - \mathbb{A}\|_{\text{op}}$ for (23a) is the first step in extending our results in Theorem 2 to vector systems. Finding an upper bound on the deception probability for vector linear and scalar nonlinear systems, where the attacker can use *any* learning algorithm, is left open for future work.

B. Model-Based Versus Model-Free

In this work, we mainly concentrated on linear systems; we assumed that the attacker constructs the fictitious sensor reading, in a model-based manner, according to the linear model (2) and its vector variants. In general, as discussed in Appendix A available online in [52], the system can be nonlinear, and the attacker might not be aware of the linearity or nonlinearity of the dynamics. Comparing the deception probability for the vast range of model-free and model-based learning methods [41], [45], [62] is an interesting research venue.

C. Continuous Learning and Hijacking

In this work, we assumed two disjoint consecutive phases (recall Section II-A): learning and hijacking, which are akin to the exploration and exploitation phases of reinforcement learning (RL) [63]. Indeed, in this two-phase process, the attacker explores the system until it reaches a desired deception probability and then moves to the exploitation phase, during which it drives the system to instability as quickly as it can. The two phases are completely separate due to the inherent tension between them: exploiting the system without properly exploring it during the learning (silent) phase increases the chances of being detected.

Despite the naturalness of two-phase attacks, just like in RL [63], one may consider more general strategies where exploration and exploitation are intertwined and gradual: as time elapses, the attacker can gain better estimates of the observed system and *gradually* increase its *detection-limited* attack. In these terms, our two-phase attack can be regarded as a two-stage approximation of the gradual attack and provides achievability bounds for such attacks. Studying more general attacks is a research venue that is left for future study.

D. Oblivious Controller

A more realistic scenario is the one in which neither the attacker nor the controller is 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 its own learning process, while, on the other hand, wants to slow down the learning process of the attacker.

In such a situation, standard adaptive control methods are clearly insufficient, as no asymmetry between the attacker and the controller can be achieved under the setting of Section II. To create a security leverage over the attacker, the controller needs to utilize a judicious privacy-enhancing signal: A properly designed privacy-enhancing signal should enjoy a positive double effect by facilitating the learning process of the controller while hindering that of the attacker at the same time. Note that, in such a scenario, while the controller knows both $\bar{U}_k$ and Γ_k, the attacker is cognizant of only their sum (17)—U_k. This is reminiscent of strategic information transfer [64].

Finally, we note that, unless the controller is able to detect an MITM attack (the attacker's hijacking phase), its learning process will be hampered by the fictitious signal that is generated according to the virtual system of the attacker (2).

E. Moving Target Defense

In our setup, the attacker has full access to the control signal (see Fig. 1), and at time $k+1$, the attacker uses the control input U_k to construct the fictitious sensor reading V_{k+1} according to (2). Thus, the *watermarking* signal [26], the private random signature, might not be an effective way to counter the learning based attacks [cf. Remark 9]. Here, we introduced the *privacy-enhancing* signal (17) to impede the learning process of the attacker and decrease the deception probability. Another technique that has been developed in the literature to counter attacks, where the attacker has full system knowledge, is having the controller covertly introduce virtual state variables unknown to the attacker but that are correlated with the ordinary state variables, so that modification of the original states will impact the extraneous states. These extraneous states can act as a *moving target* [35]–[38] for the attacker. A similar technique is the so-called *baiting*, which adds an offset to the system dynamic [39], [40]. In practice, this technique breaks the information symmetry between the attacker—which has the full system knowledge—and the controller. Using such defense techniques to hamper the learning process of our proposed attacker is an interesting research venue. In this context, the controller, by potentially sacrificing the optimally of the control task, can act in an adversarial learning setting. Assuming that the control can

covertly introduce a virtual part to the dynamics, for any given duration of the learning phase L (see Fig. 1), sufficiently fast changes in the cipher part of the dynamic can drastically hamper the learning process of the attacker. Also, as discussed in [52, Appendix A], adding a rich nonlinearity to the dynamics can be used as a way to secure the system against learning-based attacks.

F. Optimal Testing

Throughout this work, we have assumed that the controller tests the integrity of the system at a specific time step T, which 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.

In addition, following [29], we have considered the variance-based test, which searches for anomalies in the empirical variance, i.e., whether it falls outside a confidence interval of length 2δ [cf. (5)]. Studying the optimal detector for learning-based attacks is an interesting research venue.

G. Further Future Directions

Other future directions can explore the extension of the established results to partially observable linear vector systems where the input (actuation) gain is unknown, characterizing securable and unsecurable subspaces [65] for learning-based attacks, revising the attacker full access to both sensor and control signals, designing optimal privacy-enhancing signals (recall Remark 7) for linear and nonlinear systems, investigating the scenario in which the attacker is oblivious of the noise covariance matrix or more generally the noise statistics, and studying the relation between our proposed *privacy-enhancing* signal with the noise signal utilized to achieve *differential privacy* [66]. Finally, note that we assume that the attacker does not know the control strategy applied by the controller but is privy to the control actions generated by the controller. In the case where the attacker knows the control policy (or control objectives), it can do a much better job in learning the system dynamics, which is an interesting research venue.

References

[1] M. J. Khojasteh, A. Khina, M. Franceschetti, and T. Javidi, "Authentication of cyber-physical systems under learning-based attacks," *IFAC-PapersOnLine*, vol. 52, no. 20, pp. 369–374, 2019.

[2] B. Kehoe, S. Patil, P. Abbeel, and K. Goldberg, "A survey of research on cloud robotics and automation," *IEEE Trans. Autom. Sci. Eng.*, vol. 12, no. 2, pp. 398–409, Apr. 2015.

[3] D. I. Urbina *et al.*, "Limiting the impact of stealthy attacks on industrial control systems," in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur.*, 2016, pp. 1092–1105.

[4] S. M. Dibaji, M. Pirani, D. B. Flamholz, A. M. Annaswamy, K. H. Johansson, and A. Chakrabortty, "A systems and control perspective of CPS security," *Annu. Rev. Control*, vol. 47, pp. 394–411, 2019.

[5] M. Jamei *et al.*, "Micro synchrophasor-based intrusion detection in automated distribution systems: Toward critical infrastructure security," *IEEE Internet Comput.*, vol. 20, no. 5, pp. 18–27, Sep./Oct. 2016.

[6] H. Sandberg, S. Amin, and K. H. Johansson, "Cyberphysical security in networked control systems: An introduction to the issue," *IEEE Control Syst. Mag.*, vol. 35, no. 1, pp. 20–23, Feb. 2015.

[7] 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.

[8] V. Dolk, P. Tesi, C. De Persis, and W. Heemels, "Event-triggered control systems under denial-of-service attacks," *IEEE Trans. Control Netw. Syst.*, vol. 4, no. 1, pp. 93–105, Mar. 2017.

[9] Y. Shoukry *et al.*, "SMT-based observer design for cyber-physical systems under sensor attacks," *ACM Trans. Cyber-Phys. Syst.*, vol. 2, no. 1, 2018, Art. no. 5.

[10] Y. Chen, S. Kar, and J. M. Moura, "Cyber-physical attacks with control objectives," *IEEE Trans. Autom. Control*, vol. 63, no. 5, pp. 1418–1425, May 2018.

[11] D. Shi, Z. Guo, K. H. Johansson, and L. Shi, "Causality countermeasures for anomaly detection in cyber-physical systems," *IEEE Trans. Autom. Control*, vol. 63, no. 2, pp. 386–401, Feb. 2018.

[12] S. Dibaji, M. Pirani, A. Annaswamy, K. H. Johansson, and A. Chakrabortty, "Secure control of wide-area power systems: Confidentiality and integrity threats," in *Proc. IEEE Conf. Decis. Control*, Miami Beach, FL, USA, 2018, pp. 7269–7274.

[13] R. Tunga, C. Murguia, and J. Ruths, "Tuning windowed chi-squared detectors for sensor attacks," in *Proc. Amer. Control Conf.*, Milwaukee, WI, USA, 2018, pp. 1752–1757.

[14] L. Niu, J. Fu, and A. Clark, "Minimum violation control synthesis on cyber-physical systems under attacks," in *Proc. IEEE Conf. Decis. Control*, Miami Beach, FL, USA, 2018, pp. 262–269.

[15] M. S. Chong, H. Sandberg, and A. M. Teixeira, "A tutorial introduction to security and privacy for cyber-physical systems," in *Proc. Euro. Control Conf.*, 2019, pp. 968–978.

[16] I. Tomić, M. J. Breza, G. Jackson, L. Bhatia, and J. A. McCann, "Design and evaluation of jamming resilient cyber-physical systems," in *Proc. IEEE Int. Conf. Internet Things/IEEE Green Comput. Commun./IEEE Cyber, Phys., Social Comput./IEEE Smart Data*, 2018, pp. 687–694.

[17] K. Ding, X. Ren, D. E. Quevedo, S. Dey, and L. Shi, "DOS attacks on remote state estimation with asymmetric information," *IEEE Trans. Control Netw. Syst.*, vol. 6, no. 2, pp. 653–666, Jun. 2019.

[18] A. Teixeira, I. Shames, H. Sandberg, and K. H. Johansson, "A secure control framework for resource-limited adversaries," *Automatica*, vol. 51, pp. 135–148, 2015.

[19] M. Xue, S. Roy, Y. Wan, and S. K. Das, "Security and vulnerability of cyber-physical," in *Handbook on Securing Cyber-Physical Critical Infrastructure*. San Mateo, CA, USA: Morgan Kaufmann, 2012, pp. 5–30.

[20] A. Cetinkaya, H. Ishii, and T. Hayakawa, "Networked control under random and malicious packet losses," *IEEE Trans. Autom. Control*, vol. 62, no. 5, pp. 2434–2449, May 2017.

[21] P. N. Brown, H. P. Borowski, and J. R. Marden, "Security against impersonation attacks in distributed systems," *IEEE Trans. Control Netw. Syst.*, vol. 6, no. 1, pp. 440–450, Mar. 2019.

[22] Y. W. Law, T. Alpcan, and M. Palaniswami, "Security games for risk minimization in automatic generation control," *IEEE Trans. Power Syst.*, vol. 30, no. 1, pp. 223–232, Jan. 2015.

[23] I. Shames, F. Farokhi, and T. H. Summers, "Security analysis of cyber-physical systems using $\mathcal{H}_2$ norm," *IET Control Theory Appl.*, vol. 11, no. 11, pp. 1749–1755, 2017.

[24] N. Hashemi, C. Murguia, and J. Ruths, "A comparison of stealthy sensor attacks on control systems," in *Proc. Annu. Amer. Control Conf.*, 2018, pp. 973–979.

[25] R. S. Smith, "A decoupled feedback structure for covertly appropriating networked control systems," *IFAC Proc. Vol.*, vol. 44, no. 1, pp. 90–95, 2011.

[26] Y. Mo, S. Weerakkody, and B. Sinopoli, "Physical authentication of control systems: Designing watermarked control inputs to detect counterfeit sensor outputs," *IEEE Control Mag.*, vol. 35, no. 1, pp. 93–109, Feb. 2015.

[27] M. Zhu and S. Martínez, "On the performance analysis of resilient networked control systems under replay attacks," *IEEE Trans. Autom. Control*, vol. 59, no. 3, pp. 804–808, Mar. 2014.

[28] F. Miao, M. Pajic, and G. J. Pappas, "Stochastic game approach for replay attack detection," in *Proc. IEEE Conf. Decis. Control* , Florence, Italy, 2013, pp. 1854–1859.

[29] B. Satchidanandan and P. R. Kumar, "Dynamic watermarking: Active defense of networked cyber-physical systems," *Proc. IEEE*, vol. 105, no. 2, pp. 219–240, Feb. 2017.

[30] P. Hespanhol, M. Porter, R. Vasudevan, and A. Aswani, "Statistical watermarking for networked control systems," in *Proc. Amer. Control Conf.*, Milwaukee, WI, USA, 2018, pp. 5467–5472.
[31] C. Fang, Y. Qi, P. Cheng, and W. X. Zheng, "Cost-effective watermark based detector for replay attacks on cyber-physical systems," in *Proc. Asian Control Conf.*, 2017, pp. 940–945.
[32] M. Hosseini, T. Tanaka, and V. Gupta, "Designing optimal watermark signal for a stealthy attacker," in *Proc. Euro. Control Conf.*, 2016, pp. 2258–2262.
[33] A. Ferdowsi and W. Saad, "Deep learning for signal authentication and security in massive Internet-of-things systems," *IEEE Trans. Commun.*, vol. 67, no. 2, pp. 1371–1387, Feb. 2019.
[34] H. Liu, J. Yan, Y. Mo, and K. H. Johansson, "An on-line design of physical watermarks," in *Proc. IEEE Conf. Decis. Control* , Miami Beach, FL, USA, 2018, pp. 440–445.
[35] S. Weerakkody and B. Sinopoli, "Detecting integrity attacks on control systems using a moving target approach," in *Proc. IEEE Conf. Decis. Control*, Osaka, Japan, 2015, pp. 5820–5826.
[36] A. Kanellopoulos and K. G. Vamvoudakis, "A moving target defense control framework for cyber-physical systems," *IEEE Trans. Autom. Control*, vol. 65, no. 3, pp. 1029–1043, Mar. 2020.
[37] Z. Zhang, R. Deng, D. K. Yau, P. Cheng, and J. Chen, "Analysis of moving target defense against false data injection attacks on power grid," *IEEE Trans. Inf. Forensics Secur.*, vol. 15, pp. 2320–2335, 2020.
[38] P. Griffioen, S. Weerakkody, and B. Sinopoli, "An optimal design of a moving target defense for attack detection in control systems," in *Proc. Amer. Control Conf.*, Philadelphia, PA, USA, 2019, pp. 4527–4534.
[39] D. B. Flamholz, A. M. Annaswamy, and E. Lavretsky, "Baiting for defense against stealthy attacks on cyber-physical systems," in *Proc. AIAA Scitech 2019 Forum*, 2019, Art. no. 2338.
[40] A. Hoehn and P. Zhang, "Detection of covert attacks and zero dynamics attacks in cyber-physical systems," in *Proc. Amer. Control Conf.*, Boston, MA, USA, 2016, pp. 302–307.
[41] S. Dean, H. Mania, N. Matni, B. Recht, and S. Tu, "On the sample complexity of the linear quadratic regulator," *Found. Comput. Math.*, vol. 20, pp. 633–679, 2020.
[42] M. Deisenroth and C. E. Rasmussen, "PILCO: A model-based and data-efficient approach to policy search," in *Proc. Int. Conf. Mach. Learn.*, 2011, pp. 465–472.
[43] Y. Gal, R. McAllister, and C. E. Rasmussen, "Improving PILCO with Bayesian neural network dynamics models," in *Proc. Data-Efficient Mach. Learn. Workshop/Int. Conf. Mach. Learn.*, 2016, pp. 1–7.
[44] A. Rantzer, "Concentration bounds for single parameter adaptive control," in *Proc. Amer. Control Conf.*, Milwaukee, WI, USA, 2018, pp. 1862–1866.
[45] S. Tu and B. Recht, "Least-squares temporal difference learning for the linear quadratic regulator," in *Proc. Int. Conf. Mach. Learn.*, 2018, pp. 5005–5014.
[46] T. Sarkar and A. Rakhlin, "Near optimal finite time identification of arbitrary linear dynamical systems," in *Proc. Int. Conf. Mach. Learn.*, 2019, pp. 5610–5618.
[47] F. Berkenkamp, M. Turchetta, A. Schoellig, and A. Krause, "Safe model-based reinforcement learning with stability guarantees," in *31st Int. Conf. Neural Inf. Proc. Syst.*, 2017, pp. 908–918.
[48] J. F. Fisac, A. K. Akametalu, M. N. Zeilinger, S. Kaynama, J. Gillula, and C. J. Tomlin, "A general safety framework for learning-based control in uncertain robotic systems," *IEEE Trans. Autom. Control*, vol. 64, no, 7, pp. 2737–2752, Jul. 2019.
[49] M. J. Khojasteh, V. Dhiman, M. Franceschetti, and N. Atanasov, "Probabilistic safety constraints for learned high relative degree system dynamics," in *Proc. Conf. Learn. Dyn. Control*, 2020, pp. 781–792.
[50] R. Cheng, M. J. Khojasteh, A. D. Ames, and J. W. Burdick, "Safe multi-agent interaction through robust control barrier functions with learned uncertainties," in *Proc. IEEE Conf. Decis. Control*, to be published.
[51] Y. Yuan and Y. Mo, "Security in cyber-physical systems: Controller design against known-plaintext attack," in *Proc. IEEE Conf. Decis. Control*, Osaka, Japan, 2015, pp. 5814–5819.
[52] M. J. Khojasteh, A. Khina, M. Franceschetti, and T. Javidi, "Learning-based attacks in cyber-physical systems," 2018, *arXiv:1809.06023*.
[53] B. Efron and T. Hastie, *Computer Age Statistical Inference: Algorithms, Evidence, and Data Science* (ser. Institute of Mathematical Statistics Monographs). New York, NY, USA: Cambridge Univ. Press, 2016.
[54] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York, NY, USA: Wiley, 2012.
[55] R. Durrett, *Probability: Theory and Examples*. New York, NY, USA: Cambridge Univ. Press, 2010.
[56] T. L. Lai and C. Z. Wei, "Least squares estimates in stochastic regression models with applications to identification and control of dynamic systems," *Ann. Statist.*, vol. 10, pp. 154–166, 1982.
[57] J. C. Duchi and M. J. Wainwright, "Distance-based and continuum Fano inequalities with applications to statistical estimation," 2013 *arXiv:1311.2669*.
[58] M. Raginsky, "Divergence-based characterization of fundamental limitations of adaptive dynamical systems," in *Proc. Allerton Conf. Commun., Control, Comput.*, Monticello, IL, USA, 2010, pp. 107–114.
[59] D. P. Bertsekas, *Reinforcement Learning and Optimal Control*. Belmont, MA, USA: Athena Scientific, 2019.
[60] I. Ziemann and H. Sandberg, "Parameter privacy versus control performance: Fisher information regularized control," in *Proc. Amer. Control Conf.*, Denver, CO, USA, 2020, pp. 1259–1265.
[61] M. Duflo, *Random Iterative Models*, vol. 34. New York, NY, USA: Springer, 2013.
[62] V. G. Satorras, Z. Akata, and M. Welling, "Combining generative and discriminative models for hybrid inference," in *Proc. Int. Conf. Neural Inf. Proc. Syst.*, 2019, pp. 13802–13812.
[63] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, vol. 1. Cambridge, MA, USA: MIT Press, 1998.
[64] E. Akyol, C. Langbort, and T. Başar, "Privacy constrained information processing," in *Proc. IEEE Conf. Decis. Control*, Osaka, Japan, 2015, pp. 4511–4516.
[65] B. Satchidanandan and P. Kumar, "Control systems under attack: The securable and unsecurable subspaces of a linear stochastic system," in *Emerging Applications of Conttrol and Systems Theory*. New York, NY, USA: Springer, 2018, pp. 217–228.
[66] J. Cortés, G. E. Dullerud, S. Han, J. Le Ny, S. Mitra, and G. J. Pappas, "Differential privacy in control and network systems," in *Proc. IEEE Conf. Decis. Control*, Las Vegas, NV, USA, 2016, pp. 4252–4272.
[67] D. W. Marquardt, "An algorithm for least-squares estimation of nonlinear parameters," *J. Soc. Ind. Appl. Math.*, vol. 11, no. 2, pp. 431–441, 1963.
[68] J. Umlauft, T. Beckers, M. Kimmel, and S. Hirche, "Feedback linearization using Gaussian processes," in *Proc. IEEE Conf. Decis. Control*, Melbourne, VIC, Australia, 2017, pp. 5249–5255.
[69] N. Srinivas, A. Krause, S. M. Kakade, and M. W. Seeger, "Information-theoretic regret bounds for Gaussian process optimization in the bandit setting," *IEEE Trans. Inf. Theory*, vol. 58, no. 5, pp. 3250–3265, May 2012.
[70] S. Shekhar and T. Javidi *et al.*, "Gaussian process bandits with adaptive discretization," *Electron. J. Statist.*, vol. 12, no. 2, pp. 3829–3874, 2018.
[71] S. Chen *et al.*, "Approximating explicit model predictive control using constrained neural networks," in *Proc. Amer. Control Conf.*, Milwaukee, WI, USA, 2018, pp. 1520–1527.
[72] S. R. Chowdhury and A. Gopalan, "On kernelized multi-armed bandits," in *Proc. 34th Int. Conf. Mach. Learn.*, 2017, pp. 844–853.
[73] C. K. Williams and C. E. Rasmussen, *Gaussian Processes for Machine Learning*, vol. 2, no. 3. Cambridge, MA, USA: MIT Press, 2006.
[74] F. Zhang, *Matrix Theory: Basic Results and Techniques*. New York, NY, USA: Springer, 2011.

Mohammad Javad Khojasteh (Student Member, IEEE) received the double-major B.Sc. degrees in electrical engineering and in pure mathematics from the Sharif University of Technology, Tehran, Iran, in 2015, and the M.Sc. and Ph.D. degrees in electrical and computer engineering from the University of California at San Diego, La Jolla, CA, USA, in 2017 and 2019, respectively.

He is currently a Postdoctoral Scholar with the Center for Autonomous Systems and Technologies, California Institute of Technology, Pasadena, CA, USA, and a Visitor with NASA's Jet Propulsion Laboratory, Pasadena.

Anatoly Khina (Member, IEEE) received the B.Sc., M.Sc., and Ph.D. degrees in electrical engineering from Tel Aviv University, Tel Aviv, Israel, in 2006, 2010, and 2016, respectively.

He is currently a Senior Lecturer with the School of Electrical Engineering, Tel Aviv University. He was a Postdoctoral Scholar with the Department of Electrical Engineering, California Institute of Technology, Pasadena, CA, USA, from 2015 to 2018, and a Research Fellow with the Simons Institute for the Theory of Computing, University of California, Berkeley, Berkeley, CA, USA, during the Spring of 2018.

Massimo Franceschetti (Fellow, IEEE) received the Laurea (highest Hons.) degree in computer engineering from the University of Naples Federico II, Naples, Italy, in 1997, and the M.S. and Ph.D. degrees in electrical engineering from the California Institute of Technology, Pasadena, CA, USA, in 1999 and 2003, respectively.

He is currently a Professor of Electrical and Computer Engineering with the University of California at San Diego (UCSD), La Jolla, CA, USA. Before joining UCSD, he was a Postdoctoral Scholar with the University of California at Berkeley, Berkeley, CA, USA, for two years. He has held visiting positions with the Vrije Universiteit Amsterdam, Amsterdam, The Netherlands; the École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland; and the University of Trento, Trento, Italy. His research interests include physical and information-based foundations of communication and control systems.

Dr. Franceschetti received the C. H. Wilts Prize in 2003 for best doctoral thesis in electrical engineering at Caltech, the S.A. Schelkunoff Award in 2005 for best paper in the IEEE TRANSACTIONS ON ANTENNAS AND PROPAGATION, a National Science Foundation CAREER Award in 2006, an Office of Naval Research Young Investigator Award in 2007, the IEEE Communications Society Best Tutorial Paper Award in 2010, and the IEEE Control Theory Society Ruberti Young Researcher Award in 2012. He became a Guggenheim Fellow for the natural sciences (engineering) in 2019.

Tara Javidi (Member, IEEE) received the bachelor's degree in electrical engineering from the Sharif University of Technology, Tehran, Iran, in 1996, and the M.S. degrees in electrical engineering (systems) and in applied mathematics (stochastics), in 1998 and 1999, respectively, and the Ph.D. degree in electrical engineering and computer science from the University of Michigan, Ann Arbor, MI, USA, in 2002.

From 2002 to 2004, she was an Assistant Professor of Electrical Engineering at the University of Washington, Seattle, WA, USA. In 2005, she joined the University of California at San Diego (UCSD), La Jolla, CA, USA, where she is currently a Professor of Electrical and Computer Engineering. She is a Founding Co-Director of the Center for Machine-Integrated Computing and Security and a Founding Faculty Member of the Halicioglu Data Science Institute, UCSD.

Dr. Javidi is a member of the Information Theory Society Board of Governors. She was a Distinguished Lecturer of the IEEE Information Theory Society from 2017 to 2018.