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

Cross-Layer Coordinated Attacks on Cyber-Physical Systems: A LQG Game Framework with Controlled Observations

Yunhan Huang[1], Zehui Xiong[2] and Quanyan Zhu[1]

Abstract— **This work establishes a game-theoretic framework to study cross-layer coordinated attacks on cyber-physical systems (CPSs). The attacker can interfere with the physical process and launch jamming attacks on the communication channels simultaneously. At the same time, the defender can dodge the jamming by dispensing with observations. The generic framework captures a wide variety of classic attack models on CPSs. Leveraging dynamic programming techniques, we fully characterize the Subgame Perfect Equilibrium (SPE) control strategies. We also derive the SPE observation and jamming strategies and provide efficient computational methods to compute them. The results demonstrate that the physical and cyber attacks are coordinated and depend on each other.**

On the one hand, the control strategies are linear in the state estimate, and the estimate error caused by jamming attacks will induce performance degradation. On the other hand, the interactions between the attacker and the defender in the physical layer significantly impact the observation and jamming strategies. Numerical examples illustrate the interactions between the defender and the attacker through their observation and jamming strategies.

I. Introduction

Recent progress in information and communications technologies (ICT) such as the Internet of Things (IoT) and 5G high-speed cellular networks have enhanced the connectivity among physical systems and cyber systems. However, the increasing connectivity also brings with these systems heightened concern about trustworthiness. There is an urgent need for understanding security, privacy, safety, reliability, resilience, and corresponding assurance for CPSs. Due to the multi-layer and multi-stage nature of CPSs, a cross-layer cross-stage framework is a sine qua non to understand the trustworthiness of CPSs. Most existing works on the security of CPSs often focus independently on either the physical system or the cyber system. One common assumption is that adversaries can only launch one particular type of attack at a time. For example, in [1]–[4], the authors have considered DoS attacks that jam either the observation or the control signals to deteriorate the performance of the underlying system. [5] focuses only on data injection attacks on the sensors of a control system. Studies in [6] pivot purely on the replay attacks on the operator-to-actuator channel. However, in CPSs, adversaries can leverage both cyber and physical vulnerabilities to launch coordinated attacks [7]. For example, an advanced adversary can simultaneously compromise critical sensors and control units to damage targeted CPS assets.

In this work, we build a dynamic game-theoretic framework that can incorporate various attack models. The attacker conducts both physical interferences (e.g., through either direct physical intervention or cyber hacking/data injection) and jamming attacks on the observation channel. The attacker has to intelligently coordinate her/his attacks across both the physical and the cyber layers over a finite period to maximize the system degradation with minimum effort. The defender, e.g., the controller/operator, implements her/his control and has to, at each time, decide whether to observe or not. Each observation query is associated with an observation cost. The cost can capture the limited network resources, such as power, communication, and bandwidth [8]. For example, radar measurement requires megawatts of power for sensing in military applications. Thus, the defender may choose not to observe for two purposes: One is to save limited resources; the other is to dodge the jamming. The physical process is a linear dynamical system with additive white noise. The observation is partial and noisy, whose availability depends on the defender's observation decisions and the jamming policies of the attacker.

Dynamic games have long been used to capture the cross-layer multi-stage nature of CPSs and the competing nature between the system operator and the adversary [9]. At the cyber layer of the CPSs, the dynamic games are used to model the cyber kill chains of APTs that include reconnaissance, lateral movement, and command and control. These games are often built over graphical models such as computer networks [10], [11] and attack graphs [12]. At the physical layer, the dynamic games are used to describe the interactions between operational technologies (OT) and an adversary. The game-theoretic description of the threat model at the OT level guides the design of security monitoring and control strategies that aim to reduce the risks on the controlled processes and assets [13], as well as the development of resilient control mechanisms that mitigate the impact of successful attacks [14].

Our work's main contribution is the development of a cross-layer over-stage game-theoretic framework that underpins the study of CPS under coordinated simultaneous cross-layer attacks. The generic framework captures various attack models. Its connections with existing attack models will be discussed in Remark 1.

We study the SPE strategies of this dynamic game and fully characterize the SPE strategies via two dynamic programming equations. The theoretical results show that the

[1] Y. Huang and Q. Zhu are with the Department of Electrical and Computer Engineering, New York University, 370 Jay St., Brooklyn, NY. {yh.huang, qz494}@nyu.edu

[2] Z. Xiong is with the Pillar of Information Systems Technology and Design, Singapore University of Technology and Design, Singapore. zehui.xiong@ieee.org

978-94-6384-236-5 ©2021 EUCA

control strategies are linear in the state estimate. The control gain can be computed offline, independent of the observation and the jamming strategies. The effect of jamming attacks causes the estimation error and leads to significant system degradation. The capability of physical attacks also affects the jamming and observation decisions. The SPE strategies show that the defender does not observe when the observation cost surpasses system degradation caused by estimation error. When the attacker can deploy physical attacks with low costs, there is no incentive for the defender to observe even when the observation cost is zero. Otherwise, the attacker can leverage the observation to launch more accurate physical attacks. Besides, we show that the defender makes an observation even when the defender anticipates the jamming to incur a higher cost to the attacker. More results regarding the jamming and the observation decisions will be discussed in Section III-B and Section IV.

II. Problem Formulation

We consider a linear dynamical system given by

$$x_{n+1} = Ax_n + B^d u_n^d + B^a u_n^a + Cw_n, \quad 0 \le n \le N-1, \quad (1)$$

where $x_n \in \mathbb{R}^q$ is the q-dimensional state vector at time n; and u_n^d, u_n^a, w_n are vectors of dimensions less than or equal to q. Here, u_n^d is the control of the defender while u_n^a is that of the attacker. The system noise at time n is denoted by w_n. Moreover, A, B^d, B^a, and C are matrices with appropriate dimensions. The types of physical attacks can be captured by the adversarial control matrix B^a as we will explain in Remark 1. Here, we consider time-invariant system for notational simplicity, but all results in Section III can be extended to the time-varying case without much endeavor.

The associated observation system is

$$\begin{aligned} &\tilde{y}_n = Dx_n + Ev_n, \\ &y_n^d = h^d(i_n^d, i_n^a) \cdot \tilde{y}_n, \text{ and } y_n^a = h^a(i_n^d, i_n^a) \cdot \tilde{y}_n, \end{aligned} \quad (2)$$

where v and D, E are vectors and matrices with appropriate dimensions. Here, $i_n^d \in \{0,1\}$ (resp. $i_n^a \in \{0,1\}$) is the observation (resp. jamming) decision made by the defender (resp. the attacker). We call $\tilde{y}_n$ the information vector at time n. Whether or not the vector $\tilde{y}_n$ is observed by the defender and the attacker is decided by the observation decision i_n^d and the jamming decision i_n^a according to the rule $h : \{0,1\} \times \{0,1\} \to \{0,1\}$. We suppose, whenever the attacker chooses to jam the observation, the defender will receive no information. The attacker receives the same observation information as the defender does. In this case, $h^d(i_n^d, i_n^a) = h^a(i_n^d, i_n^a) = i_n^d \cdot (1 - i_n^a)$. Hence, $y_n^d = y_n^a$ and so we use y_n instead in later discussions.

We introduce the notation $X_n = \{x_0, \cdots, x_n\}$ to denote the history of state trajectory up to time n. Similarly, we define U_n^d, U_n^a, I_n^d, I_n^a, W_n, V_n, Y_n^d, and Y_n^a for u_n^d, u_n^a, i_n^d, i_n^a, w_n, v_n, y_n^d, and y_n^a, respectively. The sequences W_{N-1} and V_{N-1} are independent stochastic processes with a joint Gaussian probability distribution described by $\mathbb{E}[w_n] = \mathbb{E}[v_n] = 0$, $\mathbb{E}[w_n w_{n'}] = \Sigma_s \delta(n - n')$, and $\mathbb{E}[v_n v_{n'}] = \Sigma_o \delta(n - n')$, where $\delta(\cdot)$ is the Kronecker delta. The initial condition x_0, independent from the system noise and the observation noise, is Gaussian distributed with mean $\bar{x}_0$ and variance Σ_0.

Information: The control sequences U_{N-1}^d (resp. U_{N-1}^a) and the observation/jamming sequence I_{N-1}^d (resp. I_{N-1}^a) are to be generated by the defender (resp. attacker). At time n, the observation i_n^d is made based on the information available to the defender, which is denoted by

$$\mathcal{F}_n = \{I_{n-1}^d, I_{n-1}^a, U_{n-1}^d, U_{n-1}^a, Y_{n-1}\}, \quad n \ge 1, \quad (3)$$

with $\mathcal{F}_0 = \varnothing$. So is the jamming decision i_n^a. The controls of both defender and attacker are made based on the information available at time n after the observation, which is denoted by

$$\bar{\mathcal{F}}_n = \{\mathcal{F}_n, I_n^d, I_n^a, Y_n\}. \quad (4)$$

We assume that the state evolution equations, observation equations, noise statistics, cost functions of the controllers, and information structures of the controllers are part of common knowledge among the players. The game is hence a complete information game.

Objectives/Targets: The cost functional of the defender and the attacker, involving quadratic costs in state and their controls, as well as the observation and cost, can be written as

$$F^d(\pi^d, \pi^a) = \mathbb{E}\left[\sum_{n=0}^{N-1} c_n(x_n, u_n^d, u_n^a, i_n^d, i_n^a) + x_N' Q_N x_N\right], \quad (5)$$

where

$$\begin{aligned} c_n(x_n, u_n^d, u_n^a, i_n^d, i_n^a) =& x_n' Q_n x_n + u_n^{d\prime} R_n^d u_n^d - u_n^{a\prime} R_n^a u_n^a \\ &+ i_n^d O_n^d - i_n^a O_n^a, \end{aligned}$$

is the instantaneous cost at stage n, and we have $F^a(\pi^d, \pi^a) = -F^d(\pi^d, \pi^a)$. Here, the matrices R_n^d, R_n^a are positive definite, the matrix Q_n^d is positive semidefinite, and the scalars O_n^d, O_n^a are nonnegative for $n = 1, 2, \cdots, N-1$. Here, O_n^d represents the observation cost for the defender while O_n^a denotes the jamming cost. For any matrix M, M' indicates the transpose of M.

Strategies: $\pi^d = (\mu^d, \nu^d)$ is the strategies of the defender, where μ^d denotes the observation strategy and ν^d denotes the control strategy. The strategy of the attacker, including the jamming strategy μ^a and the control strategy ν^a, are denoted by $\pi^a = (\mu^a, \nu^a)$. Given π^d, at stage n, the control and the observation decisions of the defender are generated as $i_n^d = \mu_n^d(\mathcal{F}_n)$ and $u_n^d = \nu_n^d(\bar{\mathcal{F}}_n)$.

The defender aims to stabilize the system with minimum control effort and at the same time observe/sample economically. The attacker possesses an opposing objective, which is to undermine the defender's effort by cross-layer coordinated attacks on both the physical layer and the communication layer. Here, we consider a zero-sum game where the defender and the attacker are strictly competitive. Our results in Section III can be easily extended to a general sum setting.

Remark 1. *The framework can also capture various attack models. For example, R_n^a going to infinity means zero physical attacks. Hence, the framework specializes to optimal* *jamming attacks studied in [3], [4]; Letting $B^a = B^d$, we can model false data injection attacks in the operator-to-actuator channel [15]; With $h(i_d, i_a) = 1 - (1 - i_d)(1 - i_a)$, the framework describes pursuit-evasion type of security problems with controlled information [16], where detecting your opponent's location will expose your own location.*

III. Theoretical Results

The non-hierarchical decision making between the two players makes Nash equilibrium a natural solution concept for this game. In a Nash equilibrium, none of the players can get better off by unilateral deviation from the equilibrium.

In finite-horizon dynamic games, the characterization and the computation of the equilibrium are usually obtained by conducting backward induction, which gives rise to the concept of Subgame Perfect Nash Equilibrium (SPNE). The SPNE is a refinement of Nash equilibrium used in dynamic games with perfect information. In a perfect information game, each player is perfectly informed of the history of what has happened so far, up to the point where it is her turn to move. This game is apparently a game with perfect information. The expected cost-to-go of the defender conditioned on the information set from the beginning of time k is

$$f_k^d(\mathcal{F}_k) = \mathbb{E}\left[\sum_{n=k}^{N-1} c_n(x_n, u_n^d, u_n^a, i_n^d, i_n^a) + x_N' Q_N x_N \middle| \mathcal{F}_k\right], \tag{6}$$

for $k = 0, 1, \cdots, N-1$. The expected cost-to-go functional of the attacker is hence $f_k^a(\mathcal{F}_k) = -f_k^d(\mathcal{F}_k)$. For each stage k, the defender and the attacker, their cost-to-go functional f_k^d and f_k^a, together with the strategies for future stages $(\pi_k^d, \pi_{k+1}^d, \cdots, \pi_{N-1}^d)$ and $(\pi_k^a, \pi_{k+1}^a, \cdots, \pi_{N-1}^a)$ constitute a subgame embedded in the original game in the original game. The original game is a subgame of itself when $k = 0$.

Definition 1. *An equilibrium is an SPNE if and only if it is a Nash equilibrium in every subgame of the original game.*

An SPNE is a Nash equilibrium for the entire game since the entire game of also a subgame when $k = 0$. In this paper, we focus on studying the SPNE of the game. The complete characterization and the computation of the SPNE are conducted via two steps. The first is to characterize the SPNE control strategies from backward induction for all possible observation decision sequences. The second is to find the SPNE observation strategies based on the values under the SPNE control strategies computed in the first step.

A. Control Strategies

Suppose that we are given a sequence of observation decisions I_{N-1}^d and a sequence of jamming decisions I_{N-1}^a. Under control strategies ν^d and ν^a, the expected cost-to-go starting from time k conditioning on the information available after the observation is

$$V_k^d(\nu^d, \nu^a) = \mathbb{E}\left[\sum_{n=k}^{N-1} c_n(x_n, u_n^d, u_n^a, i_n^d, i_n^a) + x_N' Q_N x_N \middle| \bar{\mathcal{F}}_k\right]. \tag{7}$$

Define the SPNE cost-to-go value as $V_k^{d^*}(\bar{\mathcal{F}}_k) = \min_{\nu^d} V_k^d(\nu^d, \nu^{a*})$, where $\nu^a = \arg\max V_k^d(\nu^{d^*}, \nu^a)$. The complete solution of this problem requires the knowledge of 1) the SPNE control strategies at any stage 2) the SPNE expected cost of proceeding from any state at any time to the end. The main results for the SPNE control strategies are summarized in the following theorem.

Theorem 1. *For any given observation sequence I_{N-1}^d and jamming sequence I_{N-1}^a, starting from any stage $k = 0, 1, \cdots, N-1$, the SPNE cost-to-go value to the end is*

$$\begin{aligned} V_k^{d^*} =& \mathbb{E}\left[x_k' L_k x_k \middle| \bar{\mathcal{F}}_k\right] + \sum_{n=k}^{N-1} \operatorname{Tr} \Sigma_s C' L_{n+1} C \\ &+ \sum_{n=k}^{N-1} \left(\operatorname{Tr} P_n(\bar{\mathcal{F}}_n)\varphi_n + i_n^d O_n^d - i_n^a O_n^a\right), \end{aligned} \tag{8}$$

where

$$L_n = Q_n + A'\left(L_{n+1} - L_{n+1}\begin{bmatrix} B^d & B^a \end{bmatrix} M_n^{-1} \begin{bmatrix} B^{d'} \\ B^{a'} \end{bmatrix} L_{n+1}\right) A, \tag{9}$$

for $n = 1, 2 \cdots, N-1$ with $L_N = Q_N$. The matrix

$$M_n = \begin{bmatrix} R_n^d + B^{d'} L_{n+1} B^d & B^{d'} L_{n+1} B^a \\ B^{a'} L_{n+1} B^d & B^{a'} L_{n+1} B^a - R_n^a \end{bmatrix} \tag{10}$$

is invertible provided that $R_n^a > B^{a'} L_{n+1} B^a$ for $k = 0, 1, \cdots, N-1$.

At each stage n, the SPNE control strategies of the defender and the attacker take the form of the linear state feedback control laws

$$\begin{bmatrix} u_n^{d*} \\ u_n^{a*} \end{bmatrix} = -M_n^{-1} \begin{bmatrix} B^{d'} \\ B^{a'} \end{bmatrix} L_{n+1} A \hat{x}_n, \tag{11}$$

where the estimator $\hat{x}_n = \mathbb{E}\left[x_n \middle| \bar{\mathcal{F}}_n\right]$ is given by a Kalman-type linear filter [17] operating on the observation data Y_n decided by I_n^d and I_n^a. The covariance of the estimation error associated with the filter $P_n(\bar{\mathcal{F}}_n) = \mathbb{E}\left[(x_n - \hat{x}_n)(x_n - \hat{x}_n)'\right]$ can be propagated as

$$P_n(\bar{\mathcal{F}}_n) = \begin{cases} AP_{n-1}(\bar{\mathcal{F}}_{n-1})A' + C\Sigma_s C', & \text{if } h(i_n^d, i_n^a) = 0, \\ (\mathrm{Id} - G_n D)\left(AP_{n-1}(\bar{\mathcal{F}}_{n-1})A' + C\Sigma_s C'\right) \times & \\ (\mathrm{Id} - G_n D)' + G_n E \Sigma_o E' G_n', & \text{if } h(i_n^d, i_n^a) = 1, \end{cases} \tag{12}$$

where G_n can be recognized as one of the usual Kalman filter gains with

$$\begin{aligned} G_n =& \left(AP_{n-1}(\bar{\mathcal{F}}_{n-1})A' + C\Sigma_s C'\right) D' \times \\ & \left[D\left(AP_{n-1}(\bar{\mathcal{F}}_{n-1})A' + C\Sigma_s C'\right) D' + E\Sigma_o E'\right]^{-1}. \end{aligned}$$

Here, the observation effect coefficients φ_n in Equation (8) is given as

$$\varphi_n = A' L_{n+1} \begin{bmatrix} B^d & B^a \end{bmatrix} M_n^{-1} \begin{bmatrix} B^{d'} \\ B^{a'} \end{bmatrix} L_{n+1} A. \tag{13}$$

Proof. The proof is conducted by backward induction. When $k = N$, there are no control strategies involved at this stage. Hence, we have $V_N^{d^*}(\bar{\mathcal{F}}_N) = V_N^d(\bar{\mathcal{F}}_N) = \mathbb{E}\left[x_N' Q_N x_N \middle| \bar{\mathcal{F}}_N\right]$, which agrees with Equation (8). We demonstrate that Equations (8) to (13) hold when $k = N - 1$. By definition, we have

$$V_{N-1}^{d}{}^{*} = \min_{u_{N-1}^{d}} \max_{u_{N-1}^{a}} \mathbb{E}\Big[c_{N-1}(x_{N-1}, u_{N-1}^{d}, u_{N-1}^{a}, i_{N-1}^{d}, i_{N-1}^{a}) \\ + x_N' L_N x_N \Big| \bar{\mathcal{F}}_{N-1}\Big]. \tag{14}$$

Substituting $X_N = Ax_{N-1} + B^d u_{N-1}^d + B^a u_{N-1}^a + Cw_{N-1}$, carrying out the expectation, minimizing over u_{N-1}^d, and maximizing over u_{N-1}^a yield the SPNE control for this stage

$$\begin{aligned} u_{N-1}^{d}{}^{*} &= -(R_{N-1}^d + B^{d\prime} L_N B^d)^{-1} B^{d\prime} L_N (A\hat{x}_{N-1} + B^a u_{N-1}^{a}{}^{*}), \\ u_{N-1}^{a}{}^{*} &= -(B^{a\prime} L_N B^a - R_{N-1}^a)^{-1} B^{a\prime} L_N (A\hat{x}_{N-1} + B^d u_{N-1}^{d}{}^{*}). \end{aligned} \tag{15}$$

Solving the two linear equations yields Equation (11) for the case of $n = N - 1$. Substituting the SPNE control back into Equation (14) gives

$$\begin{aligned} V_{N-1}^{d}{}^{*} =& \mathbb{E}\Big[x_{N-1}' \left(Q_{N-1} + A' L_N A\right) x_{N-1} \\ & - \hat{x}_{N-1}' A' L_N \begin{bmatrix} B^d & B^a \end{bmatrix} M_{N-1}^{-1} \begin{bmatrix} B^{d\prime} \\ B^{a\prime} \end{bmatrix} L_N A \hat{x}_{N-1} \\ & + i_{N-1}^d O_{N-1}^d - i_{N-1}^a O_{N-1}^a \Big| \bar{\mathcal{F}}_{N-1}\Big] + \operatorname{Tr} \Sigma_s C' L_N C, \end{aligned} \tag{16}$$

where we have used the fact that $\mathbb{E}\left[w_{N-1}' C' L_N C w_{N-1}\right] = \operatorname{Tr} \Sigma_s C' L_{N-1} C$. The expectation in Equation (16) must still be specified for the quadratic term involving $\hat{x}_{n-1}$. For any random vector x and any appropriately dimensioned matrix M, we have the relation $\mathbb{E}\left[x' M x\right] = \bar{x}' M \bar{x} + \mathbb{E}\left[(x - \bar{x})' M (x - \bar{x})\right]$, where $\bar{x} = \mathbb{E}\left[x\right]$. Applying this relation to the quadratic term involving $\hat{x}_{n-1}$ Equation (16) yields

$$\begin{aligned} V_{N-1}^{d}{}^{*} = & \mathbb{E}\Big[x_{N-1}' L_{N-1} x_{N-1} + (x_{N-1} - \hat{x}_{N-1})' A' L_N \begin{bmatrix} B^d & B^a \end{bmatrix} \\ & \times M_{N-1}^{-1} \begin{bmatrix} B^{d\prime} \\ B^{a\prime} \end{bmatrix} L_N A (x_{N-1} - \hat{x}_{N-1}) \Big| \bar{\mathcal{F}}_{N-1}\Big] + i_{N-1}^d O_{N-1}^d \\ & - i_{N-1}^a O_{N-1}^a + \operatorname{Tr} \Sigma_s C' L_N C. \end{aligned}$$

Using the definition of $P_{N-1}(\bar{\mathcal{F}}_{N-1})$ gives

$$\begin{aligned} V_{N-1}^{d}{}^{*} = & \mathbb{E}\Big[x_{N-1}' L_{N-1} x_{N-1} \Big| \bar{\mathcal{F}}_{N-1}\Big] + i_{N-1}^d O_{N-1}^d \\ & - i_{N-1}^a O_{N-1}^a + \operatorname{Tr} \Sigma_s C' L_N C + \operatorname{Tr} P_{N-1}(\bar{\mathcal{F}}_{N-1}) \varphi_{N-1}, \end{aligned}$$

which agrees with Equation (8), and φ_{N-1} satisfies Equation (13). The propagation of $P_n(\bar{\mathcal{F}}_n)$ in Equation (12) follows the results of [18]. Thus, we have shown that Equations (8) to (13) hold for $k = N - 1$. Suppose that the claims Equations (8) to (13) hold for an arbitrary $k + 1 \leq N$. By definition of V_k^d and the tower property of conditional expectation [19], we have

$$V_k^d(\nu^d, \nu^a) = \mathbb{E}\left[c_k(x_k, u_k^d, u_k^a, i_k^d, i_k^a) + V_{k+1}^d(\bar{\mathcal{F}}_{k+1}) \Big| \bar{\mathcal{F}}_k\right].$$

An application of dynamic programming techniques yields

$$\begin{aligned} V_k^{d*}(\bar{\mathcal{F}}_k) = & \sum_{n=k+1}^{N-1} \left[\operatorname{Tr}\left(P_n(\bar{\mathcal{F}}_n)\varphi_n + \Sigma_s C' L_{n+1} C\right) + i_n^d O_n^d - i_n^a O_n^a\right] \\ & + \min_{u_k^d} \max_{u_k^a} \mathbb{E}\left[c_k(x_k, u_k^d, u_k^a, i_k^d, i_k^a) + x_{k+1}' L_{k+1} x_{k+1} \Big| \bar{\mathcal{F}}_k\right]. \end{aligned}$$

The remaining proof, which deals the minimax term, is identical to the proof for the case when $k = N-1$. Now it remains to show that M_n is invertible when $R_n^a > B^{a\prime} L_{n+1} B^a$, which is provided in Remark 2. □

The assumptions that $R_n^a > B^{a\prime} L_{n+1} B^a$ is not stringent in the setting of adversarial attacks since the cost of injecting malicious controls into the plant is usually expensive, much higher than the normal controls implemented by the defender. To guarantee the existence of a SPNE control strategy, one also needs $R_n^d + B^{d\prime} L_{n+1} B^d > 0$. For more details about the existence a SPNE control strategy, one can refer to Section 2 of [20].

Remark 2. *It is required to calculate the matrix inverse M_n^{-1}. The Schur complement of the bottom-right-corner block in the M_n matrix is the real, symmetric matrix*

$$\begin{aligned} S_B(L_{n+1}) = R_n^d + B^{d\prime} L_{n+1} B^d + B^{d\prime} L_{n+1} B^a \times \\ (R_n^a - B^{a\prime} L_{n+1} B^a)^{-1} B^{a\prime} L_{n+1} B^d, \end{aligned}$$

which is positive definite since $R_n^a > B^{a\prime} L_{n+1} B^a$, and hence invertible. Therefore, the matrix M_n^{-1} can be factored as follows:

$$M_n^{-1} = \Omega T \Omega' \tag{17}$$

with

$$\begin{aligned} \Omega' &= \begin{bmatrix} \mathrm{Id} & B^{d\prime} L_{n+1} B^a \left[R_n^a - B^{a\prime} L_{n+1} B^a\right]^{-1} \\ 0 & \mathrm{Id} \end{bmatrix}, \\ T &= \begin{bmatrix} S_B^{-1}(L_{n+1}) & 0 \\ 0 & -\left[R_n^a - B^{a\prime} L_{n+1} B^a\right]^{-1} \end{bmatrix}. \end{aligned} \tag{18}$$

This allows the defender and the attacker to compute their SPNE control laws using explicit formulae,

$$\begin{aligned} u_n^{d*} =& \nu_n^d(\bar{\mathcal{F}}_n) = -S_B^{-1}(L_{n+1}) B^{d\prime} \times \\ & \left\{\mathrm{Id} + L_{n+1} B^a \left[R_n^a - B^{a\prime} L_{n+1} B^a\right]^{-1} B^{a\prime}\right\} L_{n+1} A \hat{x}_n \\ u_n^{a*} =& \nu_a^d(\bar{\mathcal{F}}_n) = [R_n^a - B^{a\prime} L_{n+1} B^a]^{-1} B^{a\prime} \times \\ & \Bigg(\mathrm{Id} - L_{n+1} B^d S_B^{-1}(L_{n+1}) B^{d\prime} \Big\{\mathrm{Id} + L_{n+1} B^a \times \\ & \left[R_n^a - B^{a\prime} L_{n+1} B^a\right]^{-1} B^{a\prime}\Big\}\Bigg) L_{n+1} A \hat{x}_n. \end{aligned}$$

Remark 3. *If x_0 is Gaussian distributed with mean $\bar{x}_0$ and variance Σ_0, whose realization is unknown to both players but the statistics are known to both, then the following initial conditions hold in Equation* (12)*:*

$$P_0(\bar{\mathcal{F}}_0) = \\ \begin{cases} \Sigma_0, & \text{if } h(i_0^d, i_0^a) = 0, \\ \Sigma_0 - \Sigma_0 D' [D' \Sigma_0 D + E' \Sigma_o E]^{-1} D' \Sigma_0, & \text{if } h(i_0^d, i_0^a) = 1. \end{cases}$$

Given I_{N-1}^d and I_{N-1}^a, the total expected cost for the defender with the SPNE control strategies is V_0^{d*}. The cost includes the quadratic term $\mathbb{E}\left[x_0' L_0 x_0 \Big| \bar{\mathcal{F}}_0\right]$, the accumulated cost induced by system noise $\sum_{n=0}^{N-1} \operatorname{Tr} \Sigma_s C' L_{n+1} C$, the accumulated cost induced by the estimation error $\sum_{n=0}^{N-1} \operatorname{Tr} P_n(\bar{\mathcal{F}}_n) \varphi_n$, which relies heavily on the observation and the jamming decisions, as well as the accumulated costs of observing $\sum_{n=0}^{N-1} i_n^d O_n^d$ and that of jamming $\sum_{n=0}^{N-1} i_n^a O_n^a$.

Being aware of each other's strategies as shown by Equation (15), the attacker and the defender deploy their SPNE control strategies according to linear control laws based on their estimate of the state. The jamming of the defender's observation can undermine the control performance, but at the same time, it also impairs the attacking performance at the physical layer since the attacker also suffers from less information, let alone the jamming cost O_n^a. Whether the defender should observe or not also depends on multiple factors including the cost of observation O_n^d, the control performance degradation caused by the estimation error $\operatorname{Tr} P_n(\bar{\mathcal{F}}_n)\varphi_n$, and the implicit cost of offering more information to the attacker. We will shed more ligth on the observation and the jamming strategies in Section III-B.

B. The Observation and the Jamming Strategies

In Section III-A, we have demonstrated that for any given observation sequences, the expected cost-to-go of the game under the SPNE control strategies after the observation and the jamming decision have been taken at time k is $V_k^{d^*}$. In this section, we derive the the procedures for finding the SPNE observation and jamming strategies by leveraging the results we have developed in Section III-A.

Note that instead of the inner cost-to-go $V_k^d(\bar{\mathcal{F}}_k)$, the cost-to-go before the observation and the jamming decisions are made at stage k is $f_k^d(\mathcal{F}_k)$ defined in Equation (6). The strategies to be made are the observation strategy μ_n^d, the jamming strategy μ_n^a, and the control strategies ν_n^d and ν_n^a for all $n \geq k$. By the definition of $V_k^d(\bar{\mathcal{F}}_k)$ in Equation (7), the definition of f_k^d in Equation (6) and the fact that $\mathcal{F}_k \subset \bar{\mathcal{F}}_k$, we have $f_k^d(\mathcal{F}_k) = \mathbb{E}\left[V_k^d(\bar{\mathcal{F}}_k)\middle|\mathcal{F}_k\right]$. The defender aims to find both the observation strategy and the control strategy that minimize $f_k^d(\mathcal{F}_k)$ at every stage k given $\mathcal{F}_k$ while the attacker aims to do the opposite. Since the control at time stage $n \geq k$ dose not alter the information $\mathcal{F}_k$, with a slight abuse of notation, we can write $f_k^{d^*}(\mathcal{F}_k) = \min_{\pi^d}\max_{\pi^a} f_k^d(\mathcal{F}_k) = \min_{\mu^d}\max_{\mu^a}\mathbb{E}\left[V_k^{d^*}\middle|\mathcal{F}_k\right]$.

Using Equation (12), Equation (8) can be written

$$\begin{aligned}V_k^{d^*} =& \mathbb{E}\left[x_k' L_k x_k\middle|\bar{\mathcal{F}}_k\right] + \sum_{n=k}^{N-1}\operatorname{Tr}\Sigma_s C' L_{n+1} C\\ &+ \sum_{n=k}^{N-1}\operatorname{Tr}\left[Z\left(P_{n-1}(\bar{\mathcal{F}}_{n-1})\right) - h(i_n^d, i_n^a)\times\right.\\ &\left. H(P_{n-1}\left(\bar{\mathcal{F}}_{n-1})\right)\right]\varphi_n + i_n^d O_n^d - i_n^a O_n^a,\end{aligned}$$

for $k \geq 1$, where

$$\begin{aligned}Z(P) &= APA' + C\Sigma_s C',\\ H(P) &= Z(P)D'[D(Z(P))D' + E\Sigma_o E']^{-1}DZ(P).\end{aligned}$$

We define $F_k^d(\mathcal{F}_k) = \mathbb{E}\left[V_k^{d^*}|\mathcal{F}_k\right]$. Using the tower property yields

$$F_k^d(\mathcal{F}_k) = K_k^d(\mathcal{F}_k) + J_k^d(\mu^d, \mu^a, \mathcal{F}_k),$$

where

$$\begin{aligned}K_k^d(\mathcal{F}_k) &= \mathbb{E}\left[x_k' L_k x_k\middle|\mathcal{F}_k\right] + \sum_{n=k}^{N-1}\operatorname{Tr}\Sigma_s C' L_{n+1} C,\\ J_k^d(\mu^d, \mu^a, \mathcal{F}_k) &= \sum_{n=k}^{N-1}\operatorname{Tr}\left[Z\left(P_{n-1}(\bar{\mathcal{F}}_{n-1})\right) - h(i_n^d, i_n^a)\times\right.\\ &\left. H(P_{n-1}\left(\bar{\mathcal{F}}_{n-1})\right)\right]\varphi_n + i_n^d O_n^d - i_n^a O_n^a.\end{aligned}$$

Therefore, to find the SPNE observation strategies, one can only focus on $J_k^d(\mu^d, \mu^a, \mathcal{F}_k)$, which we write as $J_k^d(\mathcal{F}_k)$ for convenience in later discussion. Let us define the quantity $J_k^{d^*}(\mathcal{F}_k) = \min_{\mu^d}\max_{\mu^a} J_k^d(\mathcal{F}_k)$.

If the SPNE observation and the jamming strategies have been given for every possible $\mathcal{F}_{k+1}$, then the rule for selecting the Nash equilibrium at stage k,

$$\begin{aligned}J_k^{d^*}(\mathcal{F}_k) =& \min_{i_k^d}\max_{i_k^a}\operatorname{Tr}\left[Z\left(P_{k-1}(\bar{\mathcal{F}}_{k-1})\right) - h(i_k^d, i_k^a)\times\right.\\ &\left. H(P_{k-1}\left(\bar{\mathcal{F}}_{k-1})\right)\right]\varphi_k + i_k^d O_k^d - i_k^a O_k^a + {J_{k+1}^d}^*(\mathcal{F}_{k+1}).\end{aligned}$$

for $k = 1, 2, \cdots, N-1$ and we have ${J_N^d}^* = 0$ by definition.

Proposition 1. *Suppose that there is a sequence of SPNE observation and jamming strategies* $\{(\mu_n^{d^*}, \mu_n^{a*}), n = k+1, \cdots, N-1\}$*. Then* ${J_{k+1}^d}^*$ *can be expressed as a function* $P_k(\bar{\mathcal{F}}_k)$*. There exists no Nash equilibrium at stage* k *if*

$$\begin{aligned}O_k^a \leq& \operatorname{Tr} H(P_{k-1})\varphi_k + {J_{k+1}^d}^*(Z(P_{k-1}))\\ &- {J_{k+1}^d}^*(Z(P_{k-1}) - H(P_{k-1})),\\ O_k^d \leq& \operatorname{Tr} H(P_{k-1})\varphi_k + {J_{k+1}^d}^*(Z(P_{k-1}))\\ &- {J_{k+1}^d}^*(Z(P_{k-1}) - H(P_{k-1})),\end{aligned} \tag{19}$$

where $P_{k-1} := P_{k-1}(\bar{\mathcal{F}}_{k-1})$ *for simplicity.*

Due to space limitation, the detailed proof is provided in an unabridged version of the paper [21]. We assume in the proof that when the margin is zero, there is no incentive for both players to act. The defender and the attacker have to decide at each stage whether to observe and to attack respectively yet simultaneously. When the conditions in Equation (19) hold, there is an incentive to observe if the attacker does not jam but there always an incentive for the attacker to jam if the defender observes. Hence, there exists no Nash equilibrium in pure strategy. Now, suppose that the defender announces its observation decision first, then the attacker chooses whether to jam. That is to say at stage n the observation and the jamming strategies can be written as $\mu_n^d(\mathcal{F}_n^d)$ and $\mu_n^a(\mathcal{F}_n^a)$, where $\mathcal{F}_n^d = \mathcal{F}_n, \mathcal{F}_n^a = \mathcal{F}_n \cup I_n^d$.

Theorem 2. *Under the information structure* $\mathcal{F}_n^d$ *and* $\mathcal{F}_n^a$*, for any stage* $k \geq 1$*, there always exists a pair of Subgame Perfect Equilibrium (SPE) strategies that depend only on* P_{k-1}*. The equilibrium at stage* k *is* $(i_k^{d^*}, i_k^{a*}) = (1, 1)$ *if*

$$\begin{aligned}O_k^d <& O_k^a < \operatorname{Tr} H(P_{k-1})\varphi_k + {J_{k+1}^d}^*(Z(P_{k-1}))\\ &- {J_{k+1}^d}^*(Z(P_{k-1}) - H(P_{k-1}));\end{aligned} \tag{20}$$

The equilibrium is $(i_k^{d*}, i_k^{a*}) = (1, 0)$ *if*

$$\begin{aligned} O_k^d < \operatorname{Tr} H(P_{k-1})\varphi_k + {J_{k+1}^d}^*(Z(P_{k-1})) \\ - {J_{k+1}^d}^*(Z(P_{k-1}) - H(P_{k-1})) \leq O_k^a; \end{aligned} \tag{21}$$

and $(i_k^{d*}, i_k^{a*}) = (0, 0)$ *otherwise. Hence,* J_k^{d*} *also depends solely on* P_{k-1}.

Proof. It is easy to see the hypothesis holds at stage $N-1$. Suppose that the hypothesis is true for stage $k+1$. Following the same argument in the proof of Proposition 1, we can arrive at the same matrix described by Table I except that now the defender announces its observation decision first, then the defender reacts. In this circumstance, we have a Stackelberg game and a Stackelberg equilibrium. When the defender chooses not to observe, the best response of the attacker is not to jam since jamming generates no benefit but additional cost of jamming. This scenario gives $(i_k^d, i_k^a) = (0, 0)$ and cost-to-go $J_k^d = \operatorname{Tr} Z(P_{k-1})\varphi_k + {J_{k+1}^d}^*(Z(P_{k-1}))$. When the defender chooses to observe, the best response of the attacker is to jam if $O_k^a < \operatorname{Tr} H(P_{k-1})\varphi_k + {J_{k+1}^d}^*(Z(P_{k-1})) - {J_{k+1}^d}^*(Z(P_{k-1}) - H(P_{k-1}))$, which gives cost-to-go $J_k^d = \operatorname{Tr} Z(P_{k-1})\varphi_k + O_k^d - O_k^a + {J_{k+1}^d}^*(Z(P_{k-1}))$. The best response becomes not jamming if $O_k^a \geq \operatorname{Tr} H(P_{k-1})\varphi_k + {J_{k+1}^d}^*(Z(P_{k-1})) - {J_{k+1}^d}^*(Z(P_{k-1}) - H(P_{k-1}))$, which produces cost-to-go $\operatorname{Tr}[Z(P_{k-1}) - H(P_{k-1})] \times \varphi_k + O_k^d + {J_{k+1}^d}^*(Z(P_{k-1}) - H(P_{k-1}))$. The defender then makes appropriate observation decision that generates the least cost-to-go by anticipating the best response of the attacker. Following this logic, we obtain that the equilibrium at stage k is $(i_k^{d*}, i_k^{a*}) = (1, 1)$ if

$$\begin{aligned} O_k^d < O_k^a < \operatorname{Tr} H(P_{k-1})\varphi_k + {J_{k+1}^d}^*(Z(P_{k-1})) \\ - {J_{k+1}^d}^*(Z(P_{k-1}) - H(P_{k-1})); \end{aligned}$$

The equilibrium is $(i_k^{d*}, i_k^{a*}) = (1, 0)$ if

$$\begin{aligned} O_k^d < \operatorname{Tr} H(P_{k-1})\varphi_k + {J_{k+1}^d}^*(Z(P_{k-1})) \\ - {J_{k+1}^d}^*(Z(P_{k-1}) - H(P_{k-1})) \leq O_k^a; \end{aligned}$$

and $(i_k^{d*}, i_k^{a*}) = (0, 0)$ otherwise.

TABLE I

THE PAYOFF MATRIX FOR THE ZERO-SUM GAME BETWEEN THE DEFENDER AND THE ATTACKER AT STAGE $k \geq 1$.

J_k^d		i_k^a	
		1	0
i_k^d	1	$\operatorname{Tr} Z(P_{k-1})\varphi_k$ $+O_k^d - O_k^a$ $+{J_{k+1}^d}^*(Z(P_{k-1}))$	$\operatorname{Tr}[Z(P_{k-1}) - H(P_{k-1})]\times$ $\varphi_k + O_k^d$ $+{J_{k+1}^d}^*(Z(P_{k-1}) - H(P_{k-1}))$
	0	$\operatorname{Tr} Z(P_{k-1})\varphi_k$ $-O_k^a$ $+{J_{k+1}^d}^*(Z(P_{k-1}))$	$\operatorname{Tr} Z(P_{k-1})\varphi_k$ $+{J_{k+1}^d}^*(Z(P_{k-1}))$

Hence, the strategies at stage k depend solely on P_{k-1}. And J_k^{d*} is a function of P_{k-1}. Here, we assume that there is no incentive to act when the margin between two actions is zero. □

Even though in some circumstances, the defender will be better off if she/he can receive the observation, she/he will not observe to avoid additional cost of observation since she/he can anticipate the observation being jammed.

Corollary 1. *Suppose that at each stage* k*, the attacker announces her/his jamming decision first, then the defender reacts; i.e.,* $\mathcal{F}_k^a = \mathcal{F}_k$ *and* $\mathcal{F}_k^d = \mathcal{F}_k \cup I_k^a$*. The equilibrium at stage* k *is* $(i_k^{d*}, i_k^{a*}) = (0, 0)$ *if*

$$\begin{aligned} O_k^d \geq \operatorname{Tr} H(P_{k-1})\varphi_k + {J_{k+1}^d}^*(Z(P_{k-1})) \\ - {J_{k+1}^d}^*(Z(P_{k-1}) - H(P_{k-1})); \end{aligned}$$

the equilibrium is $(i_k^{d*}, i_k^{a*}) = (0, 1)$

$$\begin{aligned} O_k^d + O_k^d < \operatorname{Tr} H(P_{k-1})\varphi_k + {J_{k+1}^d}^*(Z(P_{k-1})) \\ - {J_{k+1}^d}^*(Z(P_{k-1}) - H(P_{k-1})); \end{aligned}$$

and $(i_k^{d*}, i_k^{a*}) = (1, 0)$ *otherwise.*

Here, the same notation J_k^{d*} has been used for games with different information structures.

Note that $P_k = Z(P_{k-1}) - H(P_{k-1})$ if the observation is made and not jammed. Since $H(P_{k-1})$ is always positive semi-definite, at stage k, if the observation is missing, i.e., $h(i_k^d, i_k^a) = 0$ and $P_k = Z(P_{k-1})$, the covariance of estimate error P_k will be larger than the covariance when there is an observation received. Here, by saying the covariance of estimate error P_k is larger than the covariance of estimate error P_k', we mean $P_k \geq P_k'$. From Equations (13) and (17), we know

$$\varphi_k = A' L_{k+1} \begin{bmatrix} B^d & B^a \end{bmatrix} \Omega T \Omega' \begin{bmatrix} B^{d\prime} \\ B^{a\prime} \end{bmatrix} L_{k+1} A,$$

where T, as we can see in Equation (18), is a block diagonal matrix with two blocks $S_B^{-1}(L_{k+1})$ and $-(R_n^a - B^{a\prime} L_{n+1} B^a)^{-1}$. Since the former block is positive definite and the later is negative definite, φ_k is neither positive semi-definite nor negative semi-definite. That means $\operatorname{Tr} P_k \varphi_k$ could be negative. The interpretation is that a larger estimate error may not be always detrimental to the defender in the presence of attacks since the observation can help the attacker inject a better control into the physical plant. Thus, even when the cost of observation $O_k^d, k = 0, 1, \cdots, N-1$ is zero, the defender will not have incentives to observe at each stage. However, when $R_k^a \to \infty$ which means that there will be no attacks on the physical plant due to high costs, φ_k becomes positive semi-definite and in this circumstance, the defender will favor a smaller covariance.

IV. NUMERICAL STUDY

It is instructive to present some numerical studies of the results in Section III. Our focus will be given to the observation and the jamming strategies. A scalar case will suffice to illustrate the interesting nature of the observation and the jamming sequences. Our results and computational approaches can be effortlessly applied to higher dimensions.

Suppose that

$$\begin{aligned}
x_{n+1} &= ax_n + u_n^d + u_n^a + w_n,\\
\tilde{y}_n &= x_n + v_n, \text{and } y_n = i_n^d(1 - i_n^a)\tilde{y}_n,\\
\Sigma_s &= 4; \Sigma_0 = 1, \Sigma_o = \sigma,\\
Q_n &= 1, R_n^d = 1, O_n^d = o^d, O_n^d = o^a \quad 0 \le n \le N-1\\
Q_N &= 8;, R_{N-1} = 10, \text{and } R_n^a = r^a, \quad 0 \le n \le N-2,
\end{aligned}$$

and $N = 30$. The unassigned parameters include the system matrix a, the cost of physical attackers r^a, the observation noise variance σ, and the observation o^d and jamming cost o^a, which are subject to change in the experiment. The computation follows a policy iteration-based algorithm [22].

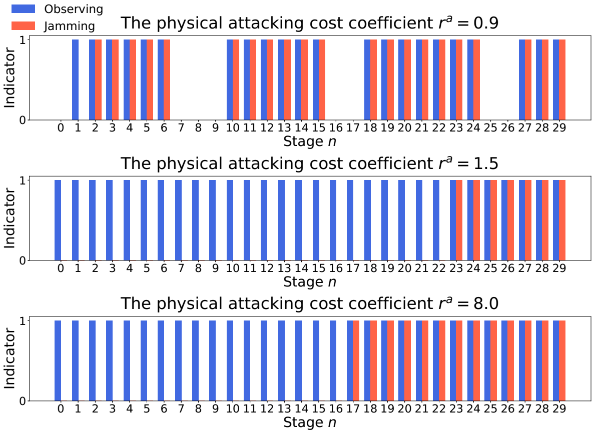

Fig. 1. The observation decisions i_n^d made by the defender and the jamming decisions i_n^a made by the attacker over 30 stages.

Figure 1 shows the observation decision sequences and the jamming sequences for various costs of physical attacks. Fixed values of $a = 0.9$, $o_d = 0$, $o_a = 15$, and $\sigma = 1$ are used. The cost of attacking the physical plant has a strong impact on both the observation and the jamming decisions. When the cost of physical attacks is low, the defender does not observe at some stages even when there is no cost of observation. This is because observations at some stages will bring additional information that can be leveraged by the attacker, who is powerful in the physical side (i.e.,low cost of attacking), to compute her/his attacks on the physical plant. As the cost of physical attacks increases (e.g. $r^a = 1.5$), undaunted by the additional information to the attacker, the defender observes at each stage. As the cost of physical further increases to $r^a = 8.0$, the attacker enjoys less benefit from additional observations since her/his physical attacks are constrained by high costs. Hence, the attacker tends to jam more to prevent the defender from receiving observations.

Figure 2 shows the observation decision sequences and the jamming sequences under different system parameters. Fixed values of $r_a = 1.5$, $o_d = 0$, $o_a = 15$, and $\sigma = 1$ are used. Sequences are shown for $a = 0.5$, a highly stable system, $a = 0.9$, a slightly stable system, and $a = 1.1$, an unstable system. In all three cases, the defender chooses to observe at every stage because physical attacks are limited by a high cost. Additional information will benefit the defender more. A highly stable system (i.e. $a = 0.5$) suffer a very low-performance degradation from missing observations. Hence, intimidated by the cost of jamming, the attacker has no incentive to jam at all. A slightly stable system however can be more easily affected by the attacker through jamming. The attacker tends to jam economically. That is to jam near the end to induce a considerable loss to the defender because $Q_N = 8$ is much higher than $Q_n = 1$ for $n \le N - 1$. Under an unstable system, the attacker simply jams every stage so that the defender cannot stabilize the system due to missing observation. The defender could have chosen not to observe because the observation will be jammed anyway. But this is a zero-sum game, so the defender can at least gain a little from the attacker's cost induced by jamming.

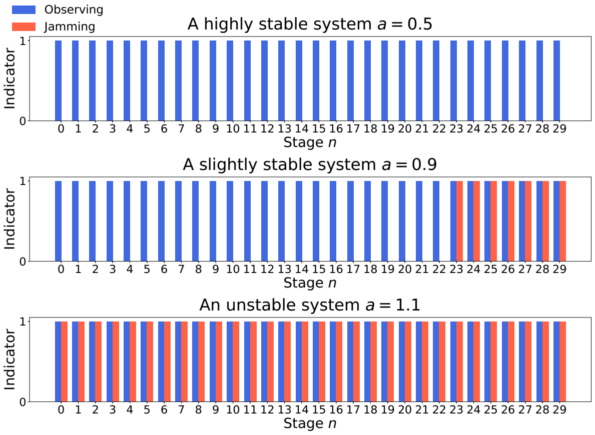

Fig. 2. The observation decisions i_n^d made by the defender and the jamming decisions i_n^a made by the attacker over 30 stages.

Figure 3 shows the number of observations and jamming as a function of observation noise variance σ. Fixed values $a = 1.1$, $o_d = 25$, $o_a = 40$, and $r_a = 20$ are used. The cost of physical attacks is set to be high which means the attacker has limited capability at the physical side. When the cost of jamming is also high $o^a = 6000$, the game resembles an optimal control problem with observations that are costly and controlled. The curve regarding the number of observations shows some unusual results. As the observation noise variance σ increases, the observation will be considered to be less valuable intuitively since it will contain less useful information about the state of the system. Grounded on this argument, one would expect the number of observations goes down monotonically to zero as σ goes to infinity. Economically, it means that we should never pay for worthless information. However, the first blue curve of Figure 3 indicates that when the observation noise variance grows from a small to a moderately large value, it is better to actually increase the number of observations. This means when the information content of each individual observation is degraded slightly, it is better to pay the cost of making extra observations in order to make a better estimate. When the cost of jamming is lower, say when $o^a = 40$, the defender observes more frequently when the observation noise variance is low because the defender knows that the attacker will be jamming (if the attacker does not jam, the defender will receive high-quality information to stabilize the system while the attacker can do little with the information). Doing so will render the attacker suffer more cost of jamming due to the fact that $O_n^a > O_n^d$. As the observation noise variance increases, the information becomes less useful. Only at some stages, the attacker has incentives to observe but most of these observations will be jammed since additional information is favored less by the attacker than the defender when the cost of physical attacks is high.

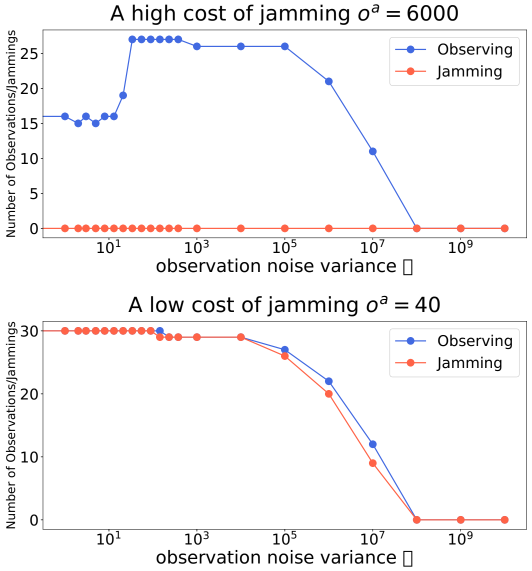

Fig. 3. The number of observations and jamming attacks in 30 stages.

V. Conclusions

In this work, we have established a cross-layer multi-stage framework to facilitate the study of CPSs under coordinated cross-layer attacks. We have demonstrated that the framework is generic and can be specified to several classic attack models. The framework has been captured by a zero-sum linear quadratic Gaussian game with controlled observation. We have built solid theoretical underpinnings for this framework which can be used to analyze a wide variety of attacking settings. The theoretical results have shown that control performance depends on the observation and jamming strategies, which affects the quality of state estimation. Hence, the observation and jamming decisions can be carried out through dynamic equations that evolve as the estimation error variance propagates. Beyond that, the capability of altering the physical process will affect the jamming and observation decisions.

Future works will focus on the study of mixed strategies of observation and jamming, investigating the continuous-time scenario, and infinite-horizon problems.

References

[1] S. Feng, A. Cetinkaya, H. Ishii, P. Tesi, and C. De Persis, "Networked control under dos attacks: Trade-offs between resilience and data rate," *IEEE Transactions on Automatic Control*, 2020.

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

[3] A. Gupta, C. Langbort, and T. Başar, "Optimal control in the presence of an intelligent jammer with limited actions," in *49th IEEE Conference on Decision and Control (CDC)*. IEEE, 2010, pp. 1096–1101.

[4] H. Zhang, P. Cheng, L. Shi, and J. Chen, "Optimal dos attack scheduling in wireless networked control system," *IEEE Transactions on Control Systems Technology*, vol. 24, no. 3, pp. 843–852, 2015.

[5] Y. Mo, R. Chabukswar, and B. Sinopoli, "Detecting integrity attacks on scada systems," *IEEE Transactions on Control Systems Technology*, vol. 22, no. 4, pp. 1396–1407, 2013.

[6] M. Zhu and S. Martinez, "On the performance analysis of resilient networked control systems under replay attacks," *IEEE Transactions on Automatic Control*, vol. 59, no. 3, pp. 804–808, 2013.

[7] S. Rass, A. Alshawish, M. A. Abid, S. Schauer, Q. Zhu, and H. De Meer, "Physical intrusion games—optimizing surveillance by simulation and game theory," *IEEE Access*, vol. 5, pp. 8394–8407, 2017.

[8] Y. Huang, V. Kavitha, and Q. Zhu, "Continuous-time markov decision processes with controlled observations," in *2019 57th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. IEEE, 2019, pp. 32–39.

[9] Y. Huang, J. Chen, L. Huang, and Q. Zhu, "Dynamic games for secure and resilient control system design," *National Science Review*, vol. 7, no. 7, pp. 1125–1141, 2020.

[10] M. A. Noureddine, A. Fawaz, W. H. Sanders, and T. Başar, "A game-theoretic approach to respond to attacker lateral movement," in *International Conference on Decision and Game Theory for Security*. Springer, 2016, pp. 294–313.

[11] K. Horák, Q. Zhu, and B. Bošanskỳ, "Manipulating adversary's belief: A dynamic game approach to deception by design for proactive network security," in *International Conference on Decision and Game Theory for Security*. Springer, 2017, pp. 273–294.

[12] T. H. Nguyen, M. Wright, M. P. Wellman, and S. Baveja, "Multi-stage attack graph security games: heuristic strategies, with empirical game-theoretic analysis," in *Proceedings of the 2017 Workshop on Moving Target Defense*, 2017, pp. 87–97.

[13] Q. Zhu and S. Rass, "On multi-phase and multi-stage game-theoretic modeling of advanced persistent threats," *IEEE Access*, vol. 6, pp. 13 958–13 971, 2018.

[14] J. Chen and Q. Zhu, "Control of multi-layer mobile autonomous systems in adversarial environments: A games-in-games approach," *IEEE Transactions on Control of Network Systems*, 2019.

[15] A. Sargolzaei, K. Yazdani, A. Abbaspour, C. D. Crane III, and W. E. Dixon, "Detection and mitigation of false data injection attacks in networked control systems," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 6, pp. 4281–4292, 2019.

[16] Y. Huang and Q. Zhu, "A pursuit-evasion differential game with strategic information acquisition," *arXiv preprint arXiv:2102.05469*, 2021.

[17] R. E. Kalman, "A New Approach to Linear Filtering and Prediction Problems," *Journal of Basic Engineering*, vol. 82, no. 1, pp. 35–45, 03 1960. [Online]. Available: https://doi.org/10.1115/1.3662552

[18] C. Cooper and N. Hahi, "An optimal stochastic control problem with observation cost," *IEEE Transactions on Automatic Control*, vol. 16, no. 2, pp. 185–189, 1971.

[19] J. Jacod and P. Protter, *Probability essentials*. Springer Science & Business Media, 2012.

[20] M. Pachter and K. Pham, "Discrete-time linear-quadratic dynamic games," *Journal of Optimization Theory and Applications*, vol. 146, no. 1, pp. 151–179, 2010.

[21] Y. Huang and Q. Zhu, "Cross-layer coordinated attacks on cyber-physical systems: A lqg game framework with controlled observations," *arXiv preprint arXiv:2012.02384*, 2020.

[22] D. P. Bertsekas and J. N. Tsitsiklis, *Neuro-dynamic programming*. Athena Scientific, 1996.