Undetectable Sensor and Actuator Attacks for Observer Based Controlled Cyber-Physical Systems

Mustafa Sinasi Ayas
Department of Electrical and Electronics Engineering
Karadeniz Technical University
61080 Trabzon, Turkey
msayas@ktu.edu.tr

Seddik M. Djouadi
Department of Electrical Engineering and Computer Science
University of Tennessee
Knoxville, TN 37996, USA
djouadi@eecs.utk.edu

Abstract— **Cyber-Physical Systems (CPSs) have vital importance because of their applications in many different areas. Attacks on these CPSs can cause considerable impact on public safety in addition to economic losses. Although studies on increasing the protection and reliability of CPSs against random malfunction are available, protection of CPSs against malignant attacks is needed. In particular, wireless sensor and actuator networks increase the attack risk. Even when a CPS functions properly, there can be undetectable attacks increasing costs or waiting for the right time to attack and destroy the CPS.**

In this paper, undetectable sensor and actuator attacks on observer-based controlled systems are theoretically analyzed. Explicit equations of both undetectable sensor and actuator signal attacks are derived. In addition, it is proved that the actuator signal attack is optimal in the sense of minimal energy attack signal. Numerical experiments are provided to validate the theoretical analyses and illustrate the effect of the undetectable attack signals.

I. INTRODUCTION

CPSs consist of physical processes, computation units and communication units which are networked through sensors, actuators, and communication devices [1]. CPSs have many applications in various fields such as robotics [2], energy [3], transportation [4], smart homes [5], health care [6], surveillance [7], and industrial process control [8]. More background about design techniques and applications of CPSs are presented in [9].

Attacks on CPSs can cause considerable damage to public safety in addition to economic losses. Safety-critical is a label used when a control application failure causes irreparable harm to people or the physical system being controlled [10]. Although the generality of cyber-attacks aimed at data networks in the past [11], there are some accomplished attacks on CPSs [12]-[15]. These successful attacks indicate that the protection mechanisms of CPSs are insufficient to assurance their healthy operation and CPSs are prone to malfunction under attacks. In addition, various studies contributed to the literature confirm this problem.

Especially the increasing demand on using wireless sensor and actuator networks, increases the attack risk on CPSs as emphasized in [3]-[8], [12]-[15]. To protect CPSs against random malfunction is studied [1]. However, protection of CPSs against malignant attacks is needed urgently, since new vulnerabilities occur as the functionalities of wireless sensor and actuator networks increase [16].

Researchers have paid attention to analyze specific attacks on network sensor and actuator data. Attacks on static state estimators by injecting false data are indicated as possible even with limited resources [17]. Deception attacks on networked control systems are presented in addition to denial of service attacks in [18]. Robust and resilient control techniques for CPSs are studied, and an application to power systems is proposed in [19]. Sensor signal attacks on observer-based controlled systems are introduced in addition to optimal sensor attack for both finite and infinite horizon linear quadratic (LQ) control in [20]. Actuator signal attacks and optimal actuator attack for both finite and infinite horizon LQ control on observer-based controlled systems are also presented in [21]. However, both undetectable sensor and undetectable actuator attacks have not been analyzed. It should be noted that even a CPS operates correctly, there can be an undetectable attack increasing costs or waiting for the right time to attack the CPS.

In this paper, undetectable sensor and actuator attacks on observer-based controlled systems are presented. Theoretical analyses are carried out in detail, and explicit equations for both of the attack signals are derived. In addition, it is proved that the actuator signal attack is optimal in the sense of minimal energy attack signal. Numerical experiments are performed in order to validate the results. These experiments are separately carried out for sensor and actuator attacks. In addition to undetectable attack signals, random impulse attack signals are also applied to the system to demonstrate the effect of the attack signals in the numerical experiments. In particular, error signals between output responses of attack free system and under attack system are illustrated.

The rest of this paper is organized as follows. Section II and Section III present undetectable sensor and actuator attacks on observer-based controlled systems, respectively. The numerical experiments including random impulse attacks and undetectable attacks for both sensor and actuator signals are given in Section IV. Finally, the conclusions and future works are drawn.

II. UNDETECTABLE SENSOR ATTACKS

An undetectable cyber-physical attack is defined as the attack that would not change the normal operating regime of a system, i.e, the output of the system under attack would be the same as normal operating regime [22]. By Lemma 4.1 [22], a sensor attack $\triangle_y(t)$ is undetectable if and only if:

$$y(x_1, u, \triangle_y(t), t) = y(x_2, u, 0, t) \tag{1}$$

where $y(\cdot)$ is the output of the system, u is control signal, and $x_1, x_2 \in \mathbb{R}^n$ are initial conditions. The attack is undetectable by static monitor for $t \in \mathbb{N}_0$, and undetectable by dynamic monitor for $t \in \mathbb{R} \geq 0$ [22]. The considered attack is undetectable by dynamic monitor in this study.

It is assumed that the attacker is able to spoof the sensor signals with a time-varying $\triangle_y(t)$ signal which starts at $t = 0$. The form of the system under attack is given in (2), where the attack signal $\triangle_y(t) \in \mathbb{R}^p$ is added to the output of the system, i. e. $y_\alpha(t) \in \mathbb{R}^p$, to spoofing the measured sensor data [20].

In order to design an observer-based controller, it is assumed that (A, B) is controllable and (A, C) is observable. Thereby, a matrix L is available such that $(A - LC)$ is Hurwitz, i.e., eigenvalues of $(A - LC)$ is in the open left half plane. Similarly, $(A - BK)$ is Hurwitz for a matrix K. The form of the state vector of the observer, i.e. $\hat{x} \in \mathbb{R}^n$, and control signal $u(t) \in \mathbb{R}^m$ are written as in (3), where $r(t) \in \mathbb{R}^m$, and $K \in \mathbb{R}^{mxn}$ are the reference input, and nominal controller, respectively.

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) \\ y_\alpha(t) &= Cx(t) + \triangle_y(t) \end{aligned} \tag{2}$$

$$\begin{aligned} \dot{\hat{x}}(t) &= A\hat{x}(t) + Bu(t) + L(y_\alpha(t) - C\hat{x}(t)) \\ u(t) &= -K\hat{x}(t) + Gr(t) \end{aligned} \tag{3}$$

By defining the error signal $e(t)$ as in (4), state-space representation of the closed-loop system given in (5), where $L \in \mathbb{R}^{nxp}$, is obtained.

$$\hat{e}(t) := x(t) - \hat{x}(t) \tag{4}$$

$$\begin{aligned} \begin{pmatrix} \dot{\hat{x}}(t) \\ \dot{\hat{e}}(t) \end{pmatrix} &= \begin{pmatrix} A - BK & LC \\ 0 & A - LC \end{pmatrix} \begin{pmatrix} \hat{x}(t) \\ \hat{e}(t) \end{pmatrix} \\ &+ \begin{pmatrix} BG \\ 0 \end{pmatrix} r(t) + \begin{pmatrix} L \\ -L \end{pmatrix} \triangle_y(t) \end{aligned} \tag{5}$$

The output of the observer-based controlled system, i. e. $y_\alpha(t) \in \mathbb{R}^p$, is:

$$y_\alpha(t) = [C \quad C] \begin{pmatrix} \dot{\hat{x}}(t) \\ \dot{\hat{e}}(t) \end{pmatrix} + \triangle_y(t) \tag{6}$$

An undetectable attack to observer-based controlled system has to satisfy the following equation written using the linearity feature of (1):

$$y_\alpha(0, u, \triangle_y, t) = y_\alpha(\begin{pmatrix} \hat{x}_0 \\ \hat{e}_0 \end{pmatrix}, u, 0, t) \tag{7}$$

Equation (7) can be rewritten using the solution of the closed-loop state-space of the system given in (5):

$$\begin{aligned} &\tilde{C} \int_0^t e^{\tilde{A}(t-\tau)} \tilde{B} u(\tau) d\tau + \tilde{C} \int_0^t e^{\tilde{A}(t-\tau)} \tilde{L} \Delta_y(\tau) d\tau \\ &+ \Delta_y(\tau) = \tilde{C} e^{\tilde{A}t} \tilde{x}_0 + \tilde{C} \int_0^t e^{\tilde{A}(t-\tau)} \tilde{B} u(\tau) d\tau, \quad t \geq 0 \end{aligned} \tag{8}$$

resulting in the undetectable sensor attack:

$$\Delta_y(\tau) = \tilde{C} e^{\tilde{A}t} \tilde{x}_0 - \tilde{C} \int_0^t e^{\tilde{A}(t-\tau)} \tilde{L} \Delta_y(\tau) d\tau, \quad t \geq 0 \tag{9}$$

where $\tilde{A} = \begin{pmatrix} A - BK & LC \\ 0 & A - LC \end{pmatrix}$, $\tilde{B} = \begin{pmatrix} BG \\ 0 \end{pmatrix}$, $\tilde{C} = [C \quad C]$, $\tilde{L} = \begin{pmatrix} L \\ -L \end{pmatrix}$, and $\tilde{x}_0 = \begin{pmatrix} \hat{x}_0 \\ \hat{e}_0 \end{pmatrix}$.

To write an explicit sensor attack equation, (9) needs some more processes. Taking the Laplace transform of (9) yields:

$$\Delta_y(s) = \tilde{C}(sI - \tilde{A})^{-1} \tilde{x}_0 - \tilde{C}(sI - \tilde{A})^{-1} \tilde{L} \Delta_y(s) \tag{10}$$

resulting in:

$$\Delta_y(s) = (I + \tilde{C}(sI - \tilde{A})^{-1} \tilde{L})^{-1} \tilde{C}(sI - \tilde{A})^{-1} \tilde{x}_0 \tag{11}$$

where $\Delta_y(s)$ is the Laplace transform of the $\Delta_y(t)$. Since $\Delta_y(s) \in L^2[0 \quad \infty]$, the Laplace transform $\Delta_y(s) \in H^2$. Equation (11) can be put in order as:

$$\Delta_y(s) = G_1 G_2 \tilde{x}_0 + \tilde{C}(sI - \tilde{A})^{-1} \tilde{x}_0 \tag{12}$$

where $G_1 = -\tilde{C}(sI - \tilde{A} + \tilde{L}\tilde{C})^{-1}$ and $G_2 = \tilde{L}\tilde{C}(sI - \tilde{A})^{-1}$. Using the notations of G_1 and G_2 given in (13), the cascaded transfer function $G_1 G_2$ is obtained [23].

$$\begin{aligned} G_1 &= \left[\begin{array}{c|c} \tilde{A} - \tilde{L}\tilde{C} & \mathrm{I} \\ \hline -\tilde{A} & 0 \end{array} \right] \\ G_2 &= \left[\begin{array}{c|c} \tilde{A} & \mathrm{I} \\ \hline \tilde{L}\tilde{C} & 0 \end{array} \right] \\ G_1 G_2 &= \left[\begin{array}{cc|c} \tilde{A} - \tilde{L}\tilde{C} & \tilde{L}\tilde{C} & 0 \\ 0 & \tilde{A} & I \\ \hline -\tilde{C} & 0 & 0 \end{array} \right] \end{aligned} \tag{13}$$

Substituting the obtained $G_1 G_2$ into (12) yields:

$$\begin{aligned} \Delta_y(s) = {} & [\tilde{C} \quad 0] \left[sI - \begin{pmatrix} \tilde{A} - \tilde{L}\tilde{C} & \tilde{L}\tilde{C} \\ 0 & \tilde{A} \end{pmatrix} \right]^{-1} \begin{bmatrix} 0 \\ I \end{bmatrix} \tilde{x}_0 \\ & + \tilde{C}(sI - \tilde{A})^{-1} \tilde{x}_0 \end{aligned} \tag{14}$$

Taking the inverse Laplace transform of (14), the explicit solution of sensor attack signal is obtained as:

$$\begin{aligned} \Delta_y(t) = {} & [\tilde{C} \quad 0] e^{\begin{pmatrix} \tilde{A} - \tilde{L}\tilde{C} & \tilde{L}\tilde{C} \\ 0 & \tilde{A} \end{pmatrix} t} \begin{bmatrix} 0 \\ I \end{bmatrix} \tilde{x}_0 \\ & + \tilde{C} e^{\tilde{A}t} \tilde{x}_0 \end{aligned} \tag{15}$$

Such an sensor attack is always possible if the attacker knows the system matrices $\tilde{A}$, $\tilde{C}$, $\tilde{L}$ and initial state $\tilde{x}_0$. This scenario reflects the case where the attacker knows or is capable of estimating the system parameters. This is the case if the adversary, for e.g. an insider or a disgruntled employee, has access to the input and output measurements of the system.

III. UNDETECTABLE ACTUATOR ATTACKS

By Lemma 4.1 [22], an actuator attack $\triangle_u(t)$ is undetectable if and only if:

$$y(x_1, u, \triangle_u(t), t) = y(x_2, u, 0, t) \tag{16}$$

where $y(\cdot)$ is the output of the system, u is control signal, and $x_1, x_2 \in \mathbb{R}^n$ are initial conditions. The considered attack is undetectable by dynamic monitor in this study.

It is assumed that the attacker is able to spoof the actuator signals with a time-varying $\triangle_u(t)$ signal which starts at $t = 0$. The form of the system under attack is given in (17), where the attack signal $\triangle_u(t) \in \mathbb{R}^m$ is added to the input of the system, i. e. $u_\alpha(t) \in \mathbb{R}^m$, to spoof the actuator signals [21].

In order to design an observer-based controller, it is assumed that (A, B) is controllable and (A, C) is observable. The form of the state vector of the observer, i.e. $\hat{x} \in \mathbb{R}^n$, and control signal $u(t) \in \mathbb{R}^m$ are written as in (18), where $r(t) \in \mathbb{R}^m$, and $K \in \mathbb{R}^{mxn}$ are the reference input, and nominal controller, respectively.

$$\begin{aligned} \dot{x}(t) &= Ax(t) + B(u_\alpha(t) + \triangle_u(t)) \\ y_\alpha(t) &= Cx(t) \end{aligned} \tag{17}$$

$$\begin{aligned} \dot{\hat{x}}(t) &= A\hat{x}(t) + B(u_\alpha(t) + \triangle_u(t)) + L(y(t) - C\hat{x}(t)) \\ u_\alpha(t) &= -K\hat{x}(t) + Gr(t) + \triangle_u(t) \end{aligned} \tag{18}$$

By defining the error signal $e(t)$ as in (4), state-space representation of the closed-loop system is obtained as:

$$\begin{aligned} \begin{pmatrix} \dot{\hat{x}}(t) \\ \dot{\hat{e}}(t) \end{pmatrix} &= \begin{pmatrix} A - BK & LC \\ 0 & A - LC \end{pmatrix} \begin{pmatrix} \hat{x}(t) \\ \hat{e}(t) \end{pmatrix} \\ &+ \begin{pmatrix} BG \\ 0 \end{pmatrix} r(t) + \begin{pmatrix} 2B \\ 0 \end{pmatrix} \triangle_u(t) \end{aligned} \tag{19}$$

where $L \in \mathbb{R}^{nxp}$. The output of the observer-based controlled system, i. e. $y_\alpha(t) \in \mathbb{R}^p$, is:

$$y_\alpha(t) = [C \quad C] \begin{pmatrix} \dot{\hat{x}}(t) \\ \dot{\hat{e}}(t) \end{pmatrix} \tag{20}$$

An undetectable attack to observer-based controlled system has to satisfy the following equation written using the linearity feature of (1):

$$y(0, u, \Delta_u, t) = y(\begin{pmatrix} \hat{x}_0 \\ \hat{e}_0 \end{pmatrix}, u, 0, t) \tag{21}$$

Equation (21) can be rewritten using the solution of the closed-loop state-space of the system given in (19):

$$\begin{aligned} \tilde{C} \int_0^t e^{\tilde{A}(t-\tau)} \tilde{B} u(\tau) d\tau + \tilde{C} \int_0^t e^{\tilde{A}(t-\tau)} \tilde{K} \Delta_u(\tau) d\tau \\ = \tilde{C} e^{\tilde{A}t} \tilde{x}_0 + \tilde{C} \int_0^t e^{\tilde{A}(t-\tau)} \tilde{B} u(\tau) d\tau, \quad t \geq 0 \end{aligned} \tag{22}$$

resulting in the undetectable actuator attack:

$$\tilde{C} \int_0^t e^{\tilde{A}(t-\tau)} \tilde{K} \Delta_u(\tau) d\tau = \tilde{C} e^{\tilde{A}t} \tilde{x}_0, \quad t \geq 0 \tag{23}$$

where $\tilde{A} = \begin{pmatrix} A - BK & LC \\ 0 & A - LC \end{pmatrix}$, $\tilde{B} = \begin{pmatrix} BG \\ 0 \end{pmatrix}$, $\tilde{C} = [C \quad C]$, $\tilde{K} = \begin{pmatrix} 2B \\ 0 \end{pmatrix}$, and $\tilde{x}_0 = \begin{pmatrix} \hat{x}_0 \\ \hat{e}_0 \end{pmatrix}$.

Since (23) is not an explicit solution more computational steps are required. The transition matrix $\phi(t, \tau)$ can be written as:

$$\phi(t, \tau) = e^{\tilde{A}(t-\tau)} = \begin{pmatrix} e^{(A-BK)(t-\tau)} & \phi_{12}(t, \tau) \\ 0 & e^{(A-LC)(t-\tau)} \end{pmatrix} \tag{24}$$

where $e^{(\cdot)t}$ is matrix exponential [24]. Differentiating the transition matrix yields:

$$\frac{\partial}{\partial t} \phi(t, 0) = \tilde{A}(t) \phi(t, 0), \quad \phi(\tau, 0) = I \tag{25}$$

$$\frac{\partial}{\partial t} \begin{pmatrix} \phi_{11} & \phi_{12} \\ 0 & \phi_{22} \end{pmatrix} = \begin{pmatrix} A - BK & LC \\ 0 & A - LC \end{pmatrix} \begin{pmatrix} \phi_{11} & \phi_{12} \\ 0 & \phi_{22} \end{pmatrix} \tag{26}$$

$$\begin{aligned} \frac{\partial}{\partial t} \phi_{12}(t, 0) &= (A - BK)\phi_{12} + LC\phi_{22}, \\ \phi_{12}(0) &= 0 \quad \phi_{22}(0) = I \end{aligned} \tag{27}$$

Therefore, $\phi_{12}(t, \tau)$ can be determined as:

$$\phi_{12}(t) = \int_0^t e^{(A-BK)(t-\tau)} LC e^{(A-LC)\tau} d\tau \tag{28}$$

The left hand side of (23) can be written as:

$$\begin{aligned} \tilde{C} \int_0^t \begin{pmatrix} e^{(A-BK)(t-\tau)} & \phi_{12}(t-\tau) \\ 0 & e^{(A-LC)(t-\tau)} \end{pmatrix} \begin{pmatrix} 2B \\ 0 \end{pmatrix} \Delta_u(\tau) d\tau \\ = 2C \int_0^t e^{(A-BK)(t-\tau)} B \Delta_u(\tau) d\tau \end{aligned} \tag{29}$$

The right hand side of (23) implies:

$$\begin{aligned} [C \quad C] \begin{pmatrix} e^{(A-BK)(t)} & \phi_{12}(t) \\ 0 & e^{(A-LC)(t)} \end{pmatrix} \begin{pmatrix} \hat{x}_0 \\ \hat{e}_0 \end{pmatrix} \\ = \begin{bmatrix} Ce^{(A-BK)(t)} & C\phi_{12}(t) + Ce^{(A-LC)t} \end{bmatrix} \begin{pmatrix} \hat{x}_0 \\ \hat{e}_0 \end{pmatrix} \end{aligned} \tag{30}$$

After all, (23) is written as:

$$\begin{aligned}2C\int_0^t e^{(A-BK)(t-\tau)}B\Delta_u(\tau)d\tau = Ce^{(A-BK)t}\hat{x}_0 \\ + \left[C\phi_{12}(t) + Ce^{(A-LC)t}\right]\hat{e}_0\end{aligned} \tag{31}$$

The explicit solution of actuator attack signal is obtained as:

$$\begin{aligned}\Delta u(\tau) = B^T e^{(A-BK)^T(t-\tau)}W^{-1}_{A-BK}(t) \\ \left(\tfrac{1}{2}e^{(A-BK)t}\hat{x}_0 + \tfrac{1}{2}\left[\phi_{12}(t) + e^{(A-LC)t}\right]\hat{e}_0\right), \quad \tau\epsilon[0,t]\end{aligned} \tag{32}$$

where W^{-1}_{A-BK} is the inverse of the controllability gramian of the pair $(A-BK, B)$. Such an actuator attack is always possible if the attacker knows the matrices A, B, C, L, K and initial state $\tilde{x}_0$. This scenario reflects the case where the attacker knows or is capable of estimating the system parameters. This is the case if the adversary, for e.g. an insider or a disgruntled employee, has access to the input and output measurements of the system.

Proof: To prove the actuator signal attack given by (32) is optimal in the sense of minimal energy attack signal, assume the undetectable attack signal is given as in (32). Then, the left hand side of (31) is:

$$\begin{aligned}2C\int_0^t e^{(A-BK)(t-\tau)}BB^T e^{(A-BK)^T(t-\tau)}d\tau W^{-1}_{A-BK}(t) \\ \left[\tfrac{1}{2}e^{(A-BK)t}\hat{x}_0 + \tfrac{1}{2}\left(\phi_{12}(t) + e^{(A-LC)t}\right)\hat{e}_0\right]\end{aligned} \tag{33}$$

The controllability gramian W_{A-BK} is:

$$\int_0^t e^{(A-BK)(t-\tau)}BB^T e^{(A-BK)^T(t-\tau)}d\tau = W_{A-BK} \tag{34}$$

Equation (35) which is equal to the right hand side of (31) is obtained by substituting (34) into (33).

$$C\left[e^{(A-BK)t}\hat{x}_0 + \left(\phi_{12}(t) + e^{(A-LC)t}\right)\hat{e}_0\right] \tag{35}$$

There are many undetectable actuator signal attack that satisfy (31), but the actuator attack signal (32) minimizes the actuator energy. To see this, define the convolution operator:

$$\Gamma : \tilde{\Delta}_u(\cdot) \mapsto 2C\int_0^t e^{(A-BK)(t-\tau)}B\tilde{\Delta}_u(\tau)d\tau \tag{36}$$

and its null space:

$$Ker(\Gamma) := \left\{\delta_u(\cdot) : 2C\int_0^t e^{(A-BK)(t-\tau)}B\delta_u(\tau)d\tau = 0 \ \right\} \tag{37}$$

Any actuator attack of the form $\tilde{\Delta}_u(t) := \Delta_u(t) + \delta_u(t)$ with $\delta_u(t) \in Ker(\Gamma)$ satisfy (32).

Then from (31):

$$\begin{aligned}2C\int_0^t e^{(A-BK)(t-\tau)}B\delta_u(\tau)d\tau = \\ 2Ce^{(A-BK)t}\int_0^t e^{-(A-BK)\tau}B\delta_u(\tau)d\tau = 0\end{aligned} \tag{38}$$

Since the system is assumed observable (38) implies that:

$$\int_0^t e^{-(A-BK)\tau}B\delta_u(\tau)d\tau = 0, \forall t \geq 0 \tag{39}$$

It follows then:

$$\begin{aligned}\left\|\tilde{\Delta}_u\right\|_2^2 &= \|\Delta_u + \delta_u\|_2^2 = \int_0^t \|\Delta_u + \delta_u\|_2^2 d\tau \\ &= \int_0^t (\Delta_u^T(\tau) + \delta_u^T(\tau))(\Delta_u(\tau) + \delta_u(\tau))d\tau \\ &= \int_0^t \Delta_u^T(\tau)\Delta_u(\tau)d\tau + \int_0^t \delta_u^T(\tau)\delta_u(\tau)d\tau + \\ &\qquad 2\int_0^t \Delta_u^T(\tau)\delta_u(\tau)d\tau \\ &= \|\Delta_u\|_2^2 + \|\delta_u\|_2^2 + 2\int_0^t \Delta_u^T(\tau)\delta_u(\tau)d\tau\end{aligned} \tag{40}$$

Note

$$\begin{aligned}\int_0^t \Delta_u^T(\tau)\delta_u(\tau)d\tau &= \int_0^t g^T(t)e^{(A-BK)(t-\tau)}B\delta_u(\tau)d\tau \\ &= g^T(t)e^{(A-BK)t}\underbrace{\int_0^t g^T(t)e^{-(A-BK)\tau}B\delta_u(\tau)d\tau}_{=0}\end{aligned} \tag{41}$$

where

$$\begin{aligned}g(t) = W^{-1}_{A-BK}(t) \\ \left[\tfrac{1}{2}e^{(A-BK)t}\hat{x}_0 + \tfrac{1}{2}\left(\phi_{12}(t) + e^{(A-LC)t}\right)\hat{e}_0\right]\end{aligned} \tag{42}$$

Therefore,

$$\begin{aligned}\left\|\tilde{\Delta}_u\right\|_2^2 &= \|\Delta_u\|_2^2 + \|\delta_u\|_2^2 \\ &\geq \|\Delta_u\|_2^2\end{aligned} \tag{43}$$

It is concluded that the actuator signal attack given by (32) is optimal in the sense of minimal energy attack signal.

IV. NUMERICAL EXPERIMENT

In order to verify the theoretical results and indicate the effectiveness of the proposed undetectable sensor and actuator attack strategies, numerical experiments implemented for both sensor and actuator attacks. In addition, random impulse sensor and actuator attack signals are applied to the observer-based controlled system to demonstrate the effectiveness of the attacks.

A system where (A,B) is controllable, and (A,C) is observable is used in the experiments. The system matrices are:

$$A=\begin{bmatrix}-2 & 1\\ 2 & -3\end{bmatrix}, B=\begin{bmatrix}1\\0\end{bmatrix}, C=\begin{bmatrix}1 & 0\end{bmatrix}$$
$$D=0, K=\begin{bmatrix}0.2915 & 0.0627\end{bmatrix}, L=\begin{bmatrix}18\\74\end{bmatrix} \tag{44}$$

A. Sensor Attacks

The closed-loop system given in (5) is considered in the sensor attack experiments. The effect of impulse sensor attack and the undetectable attack on the system output is illustrated in this section. It is assumed that the attacker is capable of accessing to the system parameters $\tilde{A}$, $\tilde{C}$, and $\tilde{L}$. The system responses are obtained utilizing the initial condition $\tilde{x}_0 = \begin{bmatrix}1 & 0 & 0 & 0\end{bmatrix}'$ in the experiments.

First, a random impulse attack signal $\Delta_y(t)$ is applied to system given in (5) to see the effect of the attack on the system response. Fig. 1 shows the system response obtained without attack and under random impulse attack with amplitude of 0.3% of the reference input. The error signal between the system responses is illustrated in Fig. 2. The failure level of the system response depends on the amplitude of the attack signal.

Then, the undetectable sensor attack signal for observer-based controlled system is calculated using (15) with the mentioned initial condition. When the calculated undetectable attack is applied to the system, the system response is obtained as illustrated in Fig. 3, where the system response without attack is also shown. The error signal between the system responses obtained without attack and under undetectable attack is given in Fig. 4. The error between the system responses are less than approximately 0.1%.

B. Actuator Attacks

In the actuator attack experiments, the closed-loop system given in (19) is considered. The effect of impulse actuator attack and the undetectable actuator attack on the system are demonstrated in this section. It is assumed that the attacker is capable of accessing to the system parameters A, B, C, L, and K. The system responses are obtained utilizing the initial condition $\tilde{x}_0 = \begin{bmatrix}1 & 0 & 0 & 0\end{bmatrix}'$ in the experiments.

In the first experiment, a random impulse attack signal $\Delta_u(t)$ is applied to system given in (19) to see the effect of the attack to the system response. Fig. 6 shows the system response obtained without attack and under random impulse attack with amplitude of 0.3% of the reference input. The error signal between the system responses is illustrated in Fig. 5. The failure level of the system response depends on the amplitude of the attack signal.

In the second experiment, the undetectable actuator attack signal for observer-based controlled system is calculated using (32) with the mentioned initial condition. When the calculated undetectable attack is applied to the system, the system response is obtained as illustrated in Fig. 7, where the system response without attack is also shown. The error signal between system responses obtained without attack and under undetectable actuator attack is given in Fig. 8. The error between the system responses are less than approximately 0.5%.

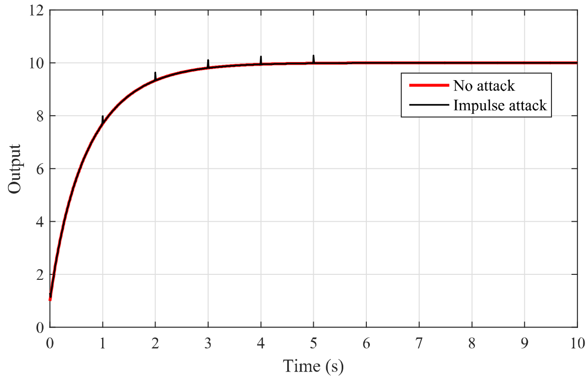

Fig. 1. System output responses obtained without attack and under random impulse sensor attack.

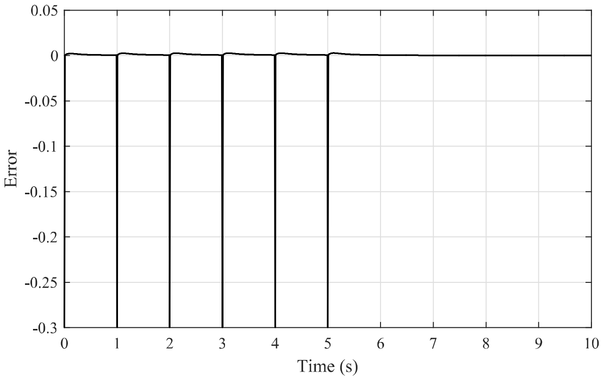

Fig. 2. The error between the system output responses obtained without attack and under random impulse sensor attack.

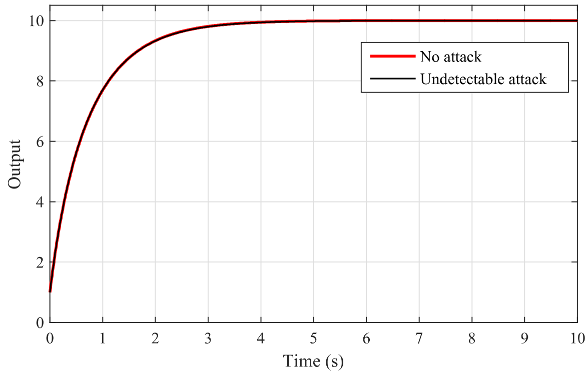

Fig. 3. System output responses obtained without attack and under undetectable sensor attack.

V. CONCLUSIONS AND FUTURE WORKS

In this study, theoretical analysis of the undetectable sensor attack and the undetectable actuator attack on observer-based control system was presented. Explicit equations of

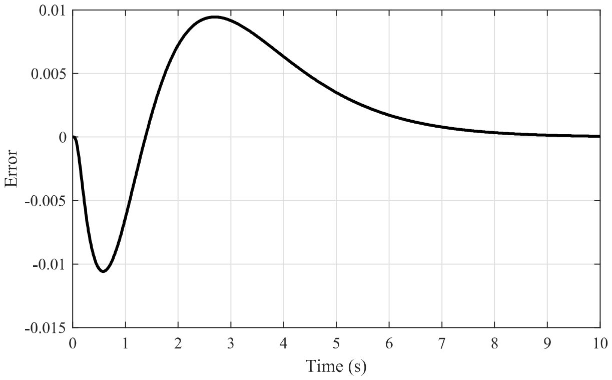

Fig. 4. The error between the system output responses obtained without attack and under undetectable sensor attack.

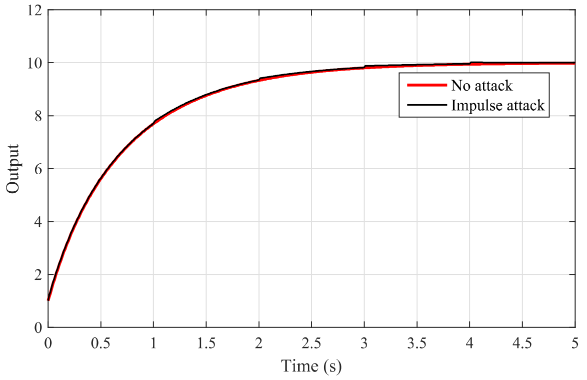

Fig. 5. System output responses obtained without attack and under undetectable actuator attack.

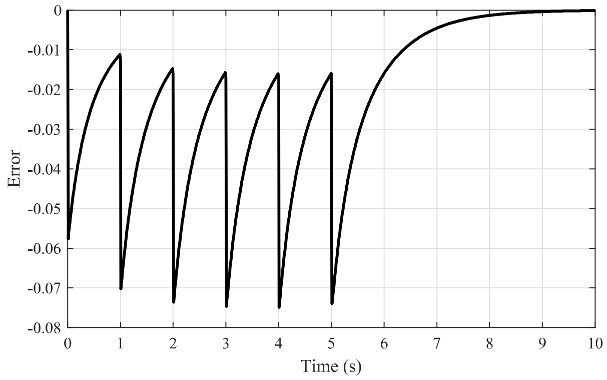

Fig. 6. The error between the system output responses obtained without attack and under random impulse actuator attack.

both undetectable sensor and actuator attack were derived. It was proved that the actuator signal attack is optimal in the sense of minimal energy attack signal.Numerical experiments were performed to validate the theoretical analyses and illustrate the effect of the attack signals. The closed-loop system responses without attack and under sensor and actuator attack were given. The error signals between output responses of attack free system and under attack signal were also illustrated. The numerical experiments indicate that

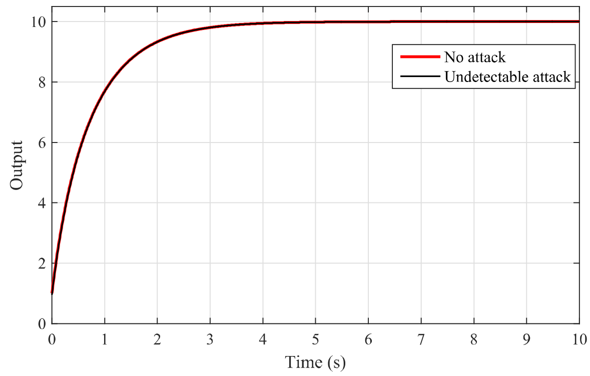

Fig. 7. System output responses obtained without attack and under undetectable actuator attack.

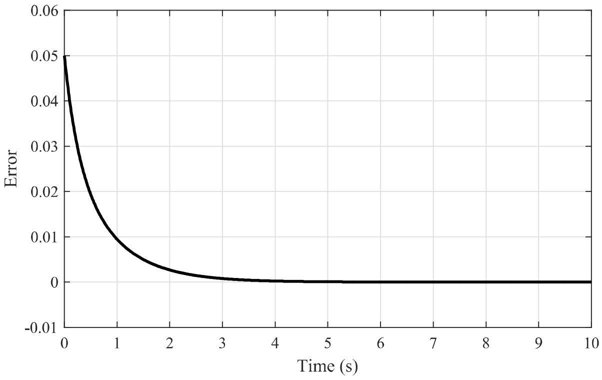

Fig. 8. The error between system output responses obtained without attack and under undetectable actuator attack.

the system response can be easily damaged by applying a random impulse attack. In addition, both undetectable sensor attack and undetectable actuator attack experiments validate the theoretical results. As a future work, detection methods for such undetectable attacks will be investigated by adding extra hardware mechanisms, and undetectable attacks on sensors and actuators for H^{∞} controllers will be derived.

VI. ACKNOWLEDGMENTS

Mustafa S. Ayas would like to thank the Scientific and Technological Research Council of Turkey (TUBITAK) BIDEB 2214 program.

REFERENCES

[1] A. A. Cardenas, S. Amin, and S. Sastry, "Secure control: Towards survivable cyber-physical systems," in *The 28th International Conference on Distributed Computing Systems Workshops*. IEEE, 2008, pp. 495–500.

[2] W. Meng, Q. Liu, W. Xu, and Z. Zhou, "A cyber-physical system for public environment perception and emergency handling," in *High Performance Computing and Communications (HPCC), 2011 IEEE 13th International Conference on*. IEEE, 2011, pp. 734–738.

[3] A. Barthels, F. Ruf, G. Walla, J. Fröschl, H.-U. Michel, and U. Baumgarten, "A model for sequence based power management in cyber physical systems," in *Information and Communication on Technology for the Fight against Global Warming*. Springer, 2011, pp. 87–101.

[4] J. K.-S. Lau, C.-K. Tham, and T. Luo, “Participatory cyber physical system in public transport application,” in *Utility and Cloud Computing (UCC), 2011 Fourth IEEE International Conference on*. IEEE, 2011, pp. 355–360.

[5] J. Kleissl and Y. Agarwal, “Cyber-physical energy systems: focus on smart buildings,” in *Proceedings of the 47th Design Automation Conference*. ACM, 2010, pp. 749–754.

[6] I. Lee and O. Sokolsky, “Medical cyber physical systems,” in *Proceedings of the 47th Design Automation Conference*. ACM, 2010, pp. 743–748.

[7] J. Chen, R. Tan, G. Xing, X. Wang, and X. Fu, “Fidelity-aware utilization control for cyber-physical surveillance systems,” *Parallel and Distributed Systems, IEEE Transactions on*, vol. 23, no. 9, pp. 1739–1751, 2012.

[8] Y. Wang, M. C. Vuran, and S. Goddard, “Cyber-physical systems in industrial process control,” *ACM Sigbed Review*, vol. 5, no. 1, p. 12, 2008.

[9] S. K. Khaitan and J. D. McCalley, “Design techniques and applications of cyberphysical systems: A survey,” *Systems Journal, IEEE*, vol. 9, no. 2, pp. 350–365, 2015.

[10] A. A. Cárdenas, S. Amin, and S. Sastry, “Research challenges for the security of control systems.” in *HotSec*, 2008.

[11] A. M. Melin, E. M. Ferragut, J. A. Laska, D. L. Fugate, and R. Kisner, “A mathematical framework for the analysis of cyber-resilient control systems,” in *Resilient Control Systems (ISRCS), 2013 6th International Symposium on*. IEEE, 2013, pp. 13–18.

[12] S. Kuvshinkova, “Sql slammer worm lessons learned for consideration by the electricity sector,” *North American Electric Reliability Council*, 2003.

[13] J. Slay and M. Miller, *Lessons learned from the maroochy water breach*. Springer, 2007.

[14] J. Conti, “The day the samba stopped,” *Engineering & Technology*, vol. 5, no. 4, pp. 46–47, 2010.

[15] N. Falliere, L. O. Murchu, and E. Chien, “W32. stuxnet dossier,” *White paper, Symantec Corp., Security Response*, vol. 5, 2011.

[16] A. Cardenas, S. Amin, B. Sinopoli, A. Giani, A. Perrig, and S. Sastry, “Challenges for securing cyber physical systems,” in *Workshop on future directions in cyber-physical systems security*, 2009, p. 5.

[17] Y. Liu, P. Ning, and M. K. Reiter, “False data injection attacks against state estimation in electric power grids,” *ACM Transactions on Information and System Security (TISSEC)*, vol. 14, no. 1, p. 13, 2011.

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

[19] Q. Zhu and T. Başar, “Robust and resilient control design for cyber-physical systems with an application to power systems,” in *Decision and Control and European Control Conference (CDC-ECC), 2011 50th IEEE Conference on*. IEEE, 2011, pp. 4066–4071.

[20] S. M. Djouadi, A. M. Melin, E. M. Ferragut, J. A. Laska, and J. Dong, “Finite energy and bounded attacks on control system sensor signals,” in *American Control Conference (ACC), 2014*. IEEE, 2014, pp. 1716–1722.

[21] S. M. Djouadi, A. M. Melin, E. M. Ferragut, J. A. Laska, J. Dong, and A. Drira, “Finite energy and bounded actuator attacks on cyber-physical systems,” in *Control Conference (ECC), 2015 European*. IEEE, 2015, pp. 3659–3664.

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

[23] K. Zhou, J. C. Doyle, K. Glover *et al.*, *Robust and optimal control*. Prentice hall New Jersey, 1996, vol. 40.

[24] W. J. Rugh, *Linear system theory*. prentice hall Upper Saddle River, NJ, 1996, vol. 2.