On Finite-State Stochastic Modeling and Secure Estimation of Cyber-Physical Systems

Dawei Shi, Robert J. Elliott, and Tongwen Chen, *Fellow, IEEE*

***Abstract*—The problem of secure state estimation and attack detection in cyber-physical systems is considered in this paper. A stochastic modeling framework is first introduced, based on which the attacked system is modeled as a finite-state hidden Markov model with switching transition probability matrices controlled by a Markov decision process. Based on this framework, a joint state and attack estimation problem is formulated and solved. Utilizing the change of probability measure approach, we show that an unnormalized joint state and attack distribution conditioned on the sensor measurement information evolves in a linear recursive form, based on which the optimal estimates can be further calculated by evaluating the normalized marginal conditional distributions. The estimation results are further applied to secure estimation of stable linear Gaussian systems, and extensions to more general systems are also discussed. The effectiveness of the results are illustrated by numerical examples and comparative simulation.**

***Index Terms*—Cyber-physical system (CPS), cyber-security, hidden Markov model (HMM), Markov decision processes, secure estimation.**

I. Introduction

THE increasing number of applications of cyber-physical systems (CPSs) have reinforced the importance of safe and secure operation of civil and industrial infrastructures [1]. Recently reported security-related accidents—e.g., the Maroochy water bleach [2], the StuxNet malware [3], and attack problems in smart grids [4]—indicate that CPSs are prone to unpredictable faults, failures, and attacks that do exist and happen during the routine operation of these systems.

Manuscript received September 6, 2015; accepted February 16, 2016. Date of publication March 14, 2016; date of current version December 26, 2016. This work was supported in part by the National Natural Science Foundation of China under Grant 61503027, in part by the Natural Sciences and Engineering Research Council of Canada, and in part by the Key Laboratory of Biomimetic Robots and Systems, Ministry of Education of China. Recommended by Associate Editor W. X. Zheng.

D. Shi is with the State Key Laboratory of Intelligent Control and Decision of Complex Systems, School of Automation, Beijing Institute of Technology, Beijing 100081, China (e-mail: dawei.shi@outlook.com).

R. J. Elliott is with School of Mathematical Sciences, University of Adelaide, Adelaide, SA 5005, Australia, and also with the Haskayne School of Business, University of Calgary, Calgary, AB T2N 1N4, Canada (e-mail: relliott@ucalgary.ca).

T. Chen is with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB T6G 1H9, Canada (e-mail: tchen@ualberta.ca).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TAC.2016.2541919

CPSs are normally composed of physical plants and complex networks of cyber-components, e.g., sensor networks, communication networks, signal processing modules. Due to the complex structure and the vulnerability of CPSs, the attacks are normally inserted into the system through the cyber-parts. Although these cyber-attacks are normally performed in a stealthy and unpredictable way, they are relatively easier to be eliminated compared with physical attacks, e.g., by changing the data encryption standard of the transmission protocol, as long as the attacks are identified. Therefore, if the attacks are detected and located in a timely fashion, the damage to the overall system can be controlled within a tolerable limit. From this perspective, attack detectors and secure estimators play a crucial role in maintaining the performance of CPSs.

During the past few years, attack detection and secure estimation have received a lot of attention in the control community. The effect of false data injection attacks was analyzed and demonstrated in [5]. For replay attacks, the feasibility condition for the attacks and countermeasures were considered in [6] and [7], and the attack detection problem was investigated in [8] utilizing a stochastic game approach. Denial-of-service attacks were studied in [9] for security-constrained optimal control and in [10] and [11] to analyze the worst-case attack policy, respectively. More recently, systematic approaches toward attack detection and secure estimation policies applicable to different attack scenarios were developed. In [12], a mathematical framework was proposed for CPSs together with attacks and monitors, and the problem of attack detection and attack set identification was solved utilizing system- and graph-theoretic approaches. In [13], the maximum number of tolerable attacks to accurately reconstruct the system state was characterized, and an efficient algorithm inspired from compressed sensing was proposed and analyzed; based on these results, a procedure for attack-resilient state estimation was proposed for systems with noise and modeling errors in [14]. In [15], a minimax optimization approach was proposed for resilient detection of a binary random state in the presence of integrity attacks. For further results and developments, see also [16]–[21] and the references therein.

In secure estimation and attack detection, it is normally assumed that the attack decision processes are completely unknown (except for the type of the attack), or that only worst-case models can be used to predict the behavior of these processes. However, for many attack policies, when the attack signals are inserted into CPSs, their effects are reflected in the available sensor measurements or at least can be measured through incorporating some special sensors designed for security monitoring purposes, as a system under attack would

normally exhibit some abnormal behavior. In this regard, it would be beneficial to associate the attack process with a fixed decision model that has a general structure allowing time-varying model parameters and access to all possible interfaces of the CPS under consideration. This would result in a number of new security-related system identification, state estimation and signal detection problems, by exploiting the information contained in the sensor measurements. In [22], a control-oriented deterministic modeling framework was introduced, in which adversary models were proposed for different attack policies.

In general, the goal of an attacker is to degrade system performance as much as possible using finite resources, which is achieved through the complex decision-making (or thinking) of the attacker and is actually the outcome of this dynamic decision procedure. In this work, together with the physical plant, the decision-making process of the attacker is modeled in a stochastic framework using Markov decision processes. This type of decision processes has been successfully utilized in modeling of complex dynamical systems, including molecular dynamics simulation [23], artificial intelligence [24], social behavior modeling [25], and human intent prediction [26]. In particular, the application of Markov decision processes in artificial intelligence [24] and human behavior modeling and understanding [25], [26] motivates us to utilize this type of processes to capture the underlying dynamic decision procedures of the attackers in CPSs, and it is the capability of the Markov processes in modeling complex decision procedures that leads to the idea of using such a model to characterize the behavior of an attacker or hacker. With this idea, an important problem to consider is the choice or estimation of the transition probabilities in Markov attack processes; to solve this problem, the joint state and attack estimation problem for prespecified transition probabilities needs to be solved first, so that efficient parameter estimation algorithms, e.g., based on expectation maximization methods [27], [28], can be developed. Alternatively, we can choose a noninformative transition probability matrix for the attack process (that is, all the elements in the matrix are equal) when we have no knowledge about the attacker, and in this regard, the only problem needs to consider is state and attack estimation. Based on the above discussions and the proposed stochastic modeling framework, a secure estimation problem is proposed and investigated in this work. The main contributions are summarized as follows.

1) A stochastic modeling framework for CPSs is proposed. The attack process is modeled by a finite-state Markov decision process and has access to both the physical plant and sensor measurement process. Based on this model, the CPS under attack is represented by a hidden Markov model with switching transition probability matrices controlled by the attack process. It is shown that the modeling framework is general enough to consider different attack policies.
2) Based on the proposed framework, a joint state and attack estimation problem is formulated and solved. We show that an unnormalized joint distribution of the state and attack signal conditioned on the measurement information evolves recursively in a linear form. The optimal state and attack estimates can be further calculated by evaluating the marginal normalized conditional distributions.
3) The application of the estimation results to stable linear Gaussian systems is considered. Numerical experience indicates that for the unattacked case, the performance of the proposed estimates obtained based on a small number of states is close to that of the Kalman filter; for the attacked case, the proposed results effectively estimate the attack signal, and the state estimation performance is satisfactory as well in the sense of average estimation error and robustness with respect to the parameters of the attack model used.

The rest of the paper is organized as follows. Section II presents the stochastic modeling framework. In Section III, the joint state and attack estimation problem is formulated and solved. Discussions on applying the results to linear Gaussian systems are provided in Section IV, where a numerical example is also presented. Concluding remarks and discussions on future work are summarized in Section V.

Notation: Let $\mathbb{N}$ denote the set of nonnegative integers. Write $\mathbb{N}_{1:M} := \{1, 2, \ldots, M\}$ and $\mathbb{Z}_{-M,N} := \{-M, -M+1, \ldots, N\}$. Let $\mathbb{I} := \{i_1, i_2, \ldots, i_N\} \subset \mathbb{N}$ be a set of indices. We use $[x_i]_{i\in\mathbb{I}}$ to denote $[x_{i_1}^\top, \ldots, x_{i_N}^\top]^\top$, where x_i can be a scalar, a vector or a matrix. For a set $\mathbb{M}$, let $|\mathbb{M}|$ be its cardinality. Let $\mathbf{1}_n$ denote $[1, 1, \ldots, 1]^\top \in \mathbb{R}^{n\times 1}$, and I_n denote the identity matrix on $\mathbb{R}^{n\times n}$. For a probability measure P (or $\overline{P}$), we use E (or $\overline{E}$) to represent the expectation operator. For a vector $v = [v_i]_{i\in\mathbb{N}_{1:n}} \in \mathbb{R}^n$, we denote $\|v\|_1$ as its 1-norm, which is defined as $\|v\|_1 = \sum_{i=1}^n \|v_i\|$, where $\|v_i\|$ is the absolute value of v_i. Let $A = [[a_{i,j}]_{j\in\mathbb{N}_{1:n}}^\top]_{i\in\mathbb{N}_{1:m}} \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{p\times q}$. Then $A \otimes B$ denotes the Kronecker product of A with B, i.e.,

$$A \otimes B := \begin{bmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{bmatrix}.$$

For matrix A, we use the right-subscript notation A_j to represent the jth column of A, and the bracketed left-subscript notation $({}_jA)$ to denote the jth row of A. Let $x, y \in \mathbb{R}^m$; then $\langle x, y\rangle := x^\top y$ denotes the inner product between x and y.

II. Stochastic Modeling Framework

For CPSs, the components which should be considered include not only the traditional components of control systems (e.g., physical plants and sensors), but also the processes that decide the potential attack signals. In this section, a stochastic modeling framework for CPSs is first introduced (see Fig. 1).

In this framework, the physical plant and the sensor measurement process are described by a finite-state hidden Markov model (HMM). The motivation of considering this model stems from the fact that HMMs represent the individual component states of a dynamic system in a natural way [29]; this type of model has been extensively utilized in speech signal processing

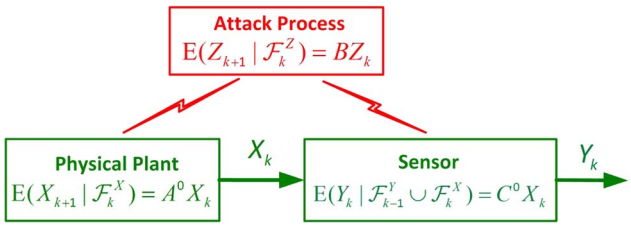


Fig. 1. Block diagram of the CPS model under attack.

(see, e.g., [30], [31] and the references therein), and has recently found applications in various areas of engineering, e.g., teleoperation systems [32], the smart grid [33], and machinery system monitoring [34]. On the other hand, considering its potential for modeling complex decision processes [23]–[26], [35], we employ a finite-state Markov decision process to model the attack process.

A. Nominal Processes

Let $(\Omega, \mathcal{F}, P)$ denote a probability space. Firstly, we introduce the nominal[1] hidden Markov model of the physical plant and the sensor measurement process on $(\Omega, \mathcal{F}, P)$. The hidden process considered is a finite-state, homogeneous, discrete-time Markov chain X. Assume the initial state X_0 is given. Suppose the cardinality of the state space of X is N. Then the state space S_X can be identified with

$$S_X = \{e_1, e_2, \ldots, e_N\}$$

where e_i is the unit vector in $\mathbb{R}^N$ with the ith element equal to 1. Let $\mathcal{F}_k^X := \sigma\{X_0, \ldots, X_k\}$, and let $\{\mathcal{F}_k^X\}$ be the complete filtration generated by $\mathcal{F}_k^X$. By the Markov property

$$P\left(X_{k+1} = e_j | \mathcal{F}_k^X\right) = P(X_{k+1} = e_j | X_k). \tag{1}$$

Let

$$a_{i,j}^0 = P(X_{k+1} = e_i | X_k = e_j)$$
$$A^0 = \left[\left[a_{i,j}^0\right]_{j\in\mathbb{N}_{1:N}}^\top\right]_{i\in\mathbb{N}_{1:N}} \in \mathbb{R}^{N\times N}. \tag{2}$$

Then

$$E\left(X_{k+1} | \mathcal{F}_k^X\right) = E(X_{k+1} | X_k) = A^0 X_k. \tag{3}$$

Let Y_k be a sensor measurement process of X_k, which takes values in a finite-state space. Let the cardinality of the state space S_Y of Y be M. Then S_Y can be identified with

$$\{f_1, f_2, \ldots, f_M\}$$

f_i being the unit vector in $\mathbb{R}^M$ with the ith element equal to 1. Write

$$C^0 = \left[\left[c_{i,j}^0\right]_{j\in\mathbb{N}_{1:N}}^\top\right]_{i\in\mathbb{N}_{1:M}}, C_j^0 = \left[c_{i,j}^0\right]_{i\in\mathbb{N}_{1:M}} \tag{4}$$

where

$$c_{i,j}^0 = P(Y_k = f_i | X_k = e_j) \tag{5}$$

[1]By "nominal" we mean that the effect of the attack signal is not considered.

so that $\sum_{i=1}^M c_{i,j}^0 = 1$ and $c_{i,j}^0 \geq 0$. Therefore

$$E(Y_k | X_k) = C^0 X_k. \tag{6}$$

Let $\mathcal{F}_k^Y$ be the completion of the σ-field on Ω generated by $Y_1, Y_2, \ldots, Y_k$. Note that

$$E(Y_k | \mathcal{F}_{k-1}^Y \cup \mathcal{F}_k^X) = E(Y_k | X_k) = C^0 X_k. \tag{7}$$

B. Process Under Attack

In the proposed framework, we consider the scenario where both the hidden process X and the measurement process Y are subject to the control of the attack process, which intentionally alters the transition matrices from A^0 and C^0 to A_k and C_k. We suppose that the pair (A_k, C_k) belongs to a finite set

$$(A_k, C_k) \in \left\{(A^1, C^1), (A^2, C^2), \ldots, (A^L, C^L)\right\}. \tag{8}$$

This implies the underlying assumption that the attacker has a finite number of choices in determining an attack policy at each time instant. Define

$$A := [A^1, A^2, \ldots, A^L] \in \mathbb{R}^{N\times(NL)} \tag{9}$$

$$C := [C^1, C^2, \ldots, C^L] \in \mathbb{R}^{M\times(NL)} \tag{10}$$

and write

$$A^l = \left[\left[a_{i,j}^l\right]_{j\in\mathbb{N}_{1:N}}^\top\right]_{i\in\mathbb{N}_{1:N}} \tag{11}$$

$$C^l = \left[\left[c_{i,j}^l\right]_{j\in\mathbb{N}_{1:N}}^\top\right]_{i\in\mathbb{N}_{1:M}}. \tag{12}$$

Considering the stealthy nature of the attack process, we assume that the actual choice (A_k, C_k) is not known, but the set of admissible modes (A^i, C^i), $i \in \mathbb{N}_{1:L}$, is known.

To model the behavior of the attacker, we assume that the attacker randomly determines the value of (A_k, C_k) according to a Markov decision process Z with

$$E\left(Z_{k+1} | \mathcal{F}_k^Z\right) = E(Z_{k+1} | Z_k) = BZ_k. \tag{14}$$

Here $Z_k \in \mathbb{R}^L$ takes its value in the set of unit vectors $\{h_i \in \mathbb{R}^L\}$, h_i being the unit vector in $\mathbb{R}^L$ with the ith element equal to 1, and

$$B = \left[\left[b_{i,j}\right]_{j\in\mathbb{N}_{1:N}}^\top\right]_{i\in\mathbb{N}_{1:L}} \in \mathbb{R}^{L\times L} \tag{14}$$

with $b_{i,j} = P(Z_{k+1} = h_i | Z_k = h_j)$. We assume that this process is independent of Y and X so that[2]

$$E\left(Z_{k+1} | \mathcal{F}_k^Y \cup \mathcal{F}_k^X \cup \mathcal{F}_k^Z\right) = E[Z_{k+1} | Z_k] = BZ_k. \tag{15}$$

If the attacker chooses (A^i, C^i), then $Z_k = h_i$. Let $\mathcal{F}_k^Z$ be the completion of the σ-field on Ω generated by $Z_1, Z_2, \ldots, Z_k$.

[2]It is possible to consider more complex and time-varying structures as well; however, a constant B matrix is utilized here to show the main idea of the modeling framework for notational simplicity.

In this way, due to the effect of unknown attacks, the attacked model of the CPS becomes

$$E\left(Y_k|\mathcal{F}_{k-1}^Y \cup \mathcal{F}_k^X \cup \mathcal{F}_k^Z\right) = E\left[Y_k|\sigma(X_k) \cup \sigma(Z_k)\right] = \sum_{i=1}^{L} C^i\langle Z_k, h_i\rangle X_k \tag{16}$$

$$E\left(X_{k+1}|\mathcal{F}_k^Y \cup \mathcal{F}_k^X \cup \mathcal{F}_k^Z\right) = E\left[X_{k+1}|\sigma(X_k) \cup \sigma(Z_k)\right] = \sum_{i=1}^{L} A^i\langle Z_k, h_i\rangle X_k. \tag{17}$$

Equations (15)–(17) describe the attacked process model under the original probability measure P. Note that in (16) and (17), the inner product $\langle Z_k, h_i\rangle$ is used as an indicator function to reflect the impact of the attack process Z_k on the probability transition matrices of X_k and Y_k, since Z_k takes value from $\{h_1, \ldots, h_L\}$, which is the set of unit vectors on $\mathbb{R}^L$.

For the proposed model, the attacks are conducted by changing the transition probabilities of the hidden process X and the measurement process Y. In particular, we note that the change of the transition matrices from A^0 and C^0 to A^i and C^i naturally models the action of changing the values of the realizations of the hidden process X and the measurement process Y from their original values to some fake values, which is what the attackers normally do during an attack event. The modeling framework introduced above can be used to consider a variety of attack policies, under the same model structure in (15)–(17). For instance, denial-of-service attacks can be characterized by introducing a "packet lost" state in the state space of the measurement process Y, and the attack policy can be modeled by appropriately choosing the relevant parameters in C^i and B; replay attacks can be described by considering a set of C^i matrices with all elements in the ith row equal to 1 and devising a particular time-varying B matrix; false data injection attacks can be modeled naturally by the time-varying and switching structure of the transition probability matrices (see Example 2 for details).

Before continuing, we first present two examples of the model introduced.

Example 1 (HMM Subject to Shifting Attacks): Consider a hidden model X with state space $\{e_i\}$ associated with a measurement process Y with state space $\{f_i\}$ and nominal transition probabilities given in (3) and (7), respectively. The policy of the attacker is to shift the state of X at time k by $\kappa_k^X \in \{1, 2, \ldots, N\}$ units and the state of Y by $\kappa_k^Y \in \{1, 2, \ldots, M\}$ units in the direction shown in Fig. 2. In this way, the attacked process has $L = MN$ modes, and each mode $Z_k = h_i$ or equivalently $(A_k, C_k) = (A^i, C^i)$, corresponds to a specific choice of combination (κ_k^X, κ_k^Y). Specifically, (A^i, C^i) satisfies

$$A^i = T_A^i A^0,\ C^i = T_C^i C^0 \tag{18}$$

where T_A^i and T_C^i are permutation matrices with appropriate sizes. To further relate the theoretic model introduced with a practical example, we note that the finite-state HMM modeling framework discussed here can be utilized to capture the scenario of machine health monitoring [36] subject to deception attacks. In this scenario, X_k represents the number of machines working in abnormal state such that $X_k = e_i$ means that the number of machines working in abnormal state is equal to i at time instant k (with N being the total number of machines), while Y_k represents the number of nonconforming products such that $Y_k = f_j$ corresponds to the event that the number of nonconforming products equals j at time instant k (with M being the total number of products). The nominal transition probability matrices A^0 and C^0 can be calculated according to the modeling procedure introduced in [36]. In machine health monitoring, the information about the number of nonconforming products is utilized by a monitor to estimate the number of abnormal machines. In this case, the action of an attacker can be to mess up the measurement information by, for instance, hacking the computer that performs the task of product quality testing and changing the measured number of nonconforming products to some fake values according to the shifting law introduced in this example. Note that in this case only the measurement process Y is attacked so that $T_A^i = I_N$ holds for all i.

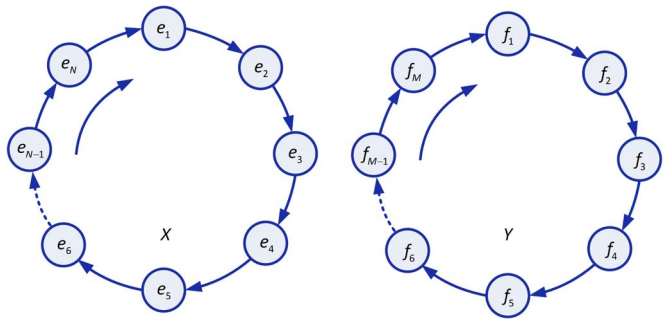


Fig. 2. HMM under shifting attacks.

Example 2: Consider a nonlinear stochastic process[3]

$$x_{k+1} = f(x_k, w_k) \tag{19}$$

where $\{x_k\}$ is assumed to be a stationary process with distribution μ_x, and $\{w_k\}$ is the input noise sequence, which is an i.i.d. stationary process with distribution μ_w. We quantize the support of x into a total of N regions, which are denoted by

$$\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N\}$$

and use a finite-state process $\{X_k\}$ taking values in $\{e_1, e_2, \ldots, e_N\}$ to represent the effect of quantization. We associate each region $\mathbf{x}_i$ with e_i, i.e.,

$$x_k \in \mathbf{x}_i \Longleftrightarrow X_k = e_i.$$

Since w_k is an i.i.d. process, $\{x_k\}$ is a Markov process, and thus $\{X_k\}$ is Markov as well. In particular, the transition probability

[3]In this example, we do not explicitly specify the dimensions of the vector-valued processes considered, as the dimension parameters do not affect our analysis; we assume that the processes are of compatible dimensions.

matrix introduced in (2) is calculated as

$$\begin{aligned} a^0_{i,j} &= P(X_{k+1} = e_i | X_k = e_j) \\ &= \int_{\mathbf{x}_j} P(x_{k+1} \in \mathbf{x}_i | x_k = x) f(x_k = x | x_k \in \mathbf{x}_j) dx \\ &= \int_{\mathbf{x}_j} P\left(w_k \in \{w | f(x,w) \in \mathbf{x}_i\}\right) \frac{\mu_x(x)}{\int_{\mathbf{x}_j} \mu_x(\xi) d\xi} dx \\ &= \int_{\mathbf{x}_j} \int_{f(x,\varrho)\in\mathbf{x}_i} \mu_w(\varrho) d\varrho \frac{\mu_x(x)}{\int_{\mathbf{x}_j} \mu_x(\xi) d\xi} dx. \end{aligned} \tag{20}$$

Note that in many situations, the above expression can be calculated offline. We suppose the state x_k is measured by a sensor through a measurement process $\{y_k\}$

$$y_k = h(x_k, v_k, z_k). \tag{21}$$

Here $\{v_k\}$ is a stationary i.i.d. measurement noise process with distribution μ_v, and z_k denotes the attack signal inserted to the sensor. We assume that the attacker determines the value of z_k from a finite set $\{\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_L\}$ according to a finite-state Markov chain Z_k given in (15). We also quantize y_k over its support, and the quantization regions are denoted as $\{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_M\}$. The processes $\{Y_k\}$ and $\{Z_k\}$ are defined as

$$\begin{aligned} y_k \in \mathbf{y}_i &\Longleftrightarrow Y_k = f_i \\ z_k = \mathbf{z}_i &\Longleftrightarrow Z_k = h_i. \end{aligned}$$

Note that conditioned on z_k and x_k, y_k does not depend on the past information of x and z. In particular, the transition probability matrix introduced in (12) is calculated as

$$\begin{aligned} c^l_{i,j} &= P(Y_k = f_i | X_k = e_j, Z_k = h_l) \\ &= P(y_k \in \mathbf{y}_i | x_k \in \mathbf{x}_j, Z_k = h_l) \\ &= \int_{\mathbf{x}_j} P(y_k \in \mathbf{y}_i | x_k = x, Z_k = h_l) \\ &\quad \cdot f(x_k = x | x_k \in \mathbf{x}_j, Z_k = h_l) dx \\ &= \int_{\mathbf{x}_j} P(y_k \in \mathbf{y}_i | x_k = x, z_k = \mathbf{z}_l) f(x_k = x | x_k \in \mathbf{x}_j) dx \\ &= \int_{\mathbf{x}_j} P\left(h(x, v_k, \mathbf{z}_l) \in \mathbf{y}_i\right) \frac{\mu_x(x)}{\int_{\mathbf{x}_j} \mu_x(\xi) d\xi} dx \\ &= \int_{\mathbf{x}_j} \int_{h(x,\varrho,\mathbf{z}_l)\in\mathbf{y}_i} \mu_v(\varrho) d\varrho \frac{\mu_x(x)}{\int_{\mathbf{x}_j} \mu_x(\xi) d\xi} dx. \end{aligned} \tag{22}$$

As a result, the quantized nonlinear stochastic process with attacks on the measurement process (e.g., certain false data injection attacks) can be recast into the set of models given in (15) –(17). Note that in this example, we have $A^l = A^0$ for all $l \in \mathbb{N}_{1:L}$.

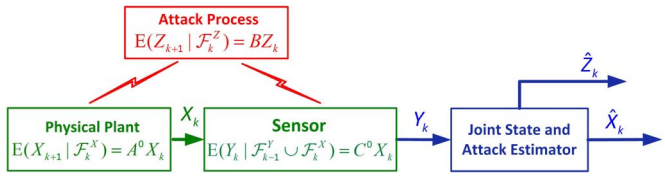


Fig. 3. Block diagram of the CPS model and the joint state and attack estimation scheme.

III. Secure Estimation

Based on the proposed modeling framework, a secure state estimation and attack detection problem is now considered (see Fig. 3). The objective is to estimate the hidden process X, as well as the attack signal Z, which is given in terms of the joint distribution of X_k and Z_k conditioned on $\mathcal{Y}_k$. Then, the estimate of X_k can be further obtained by calculating the marginal distribution of X_k conditioned on $\mathcal{Y}_k$. The exploration of this problem not only provides results and algorithms on state and attack estimation for CPSs with prior model parameters, but also serves as a starting point and a necessary step for expectation maximization solutions [37] to the problems of estimating the model parameters of the attack processes—these parameters are not normally known and thus need to be evaluated and updated online.

A. Change of Probability Measure

In this work, the secure estimation problem is investigated using the change of probability measure approach [38]. To do this, a new probability measure is firstly proposed and its properties are investigated, based on which the estimation problem is further solved.

Therefore, we introduce a new probability measure $\overline{P}$ and describe the processes under this measure. Assume under $\overline{P}$ we have

$$\begin{aligned} \overline{E}\left(X_{k+1} | F_k^Y \cup \mathcal{F}_k^X \cup \mathcal{F}_k^Z\right) &= \overline{E}\left[X_{k+1} | \sigma(X_k) \cup \sigma(Z_k)\right] \\ &= \sum_{i=1}^{L} A^i \langle Z_k, h_i \rangle X_k \end{aligned} \tag{23}$$

$$\overline{E}(Z_{k+1} | \mathcal{F}_k^Y \cup \mathcal{F}_k^X \cup \mathcal{F}_k^Z) = \overline{E}[Z_{k+1} | Z_k] = BZ_k \tag{24}$$

indicating the dynamics of X and Z are unchanged under the new measure. However, under $\overline{P}$, Y is a process independent of X and Z, and is a sequence of uniformly distributed random variables satisfying

$$\overline{E}\left(\langle Y_k, f_i \rangle | \mathcal{F}_{k-1}^Y \cup \mathcal{F}_k^X \cup \mathcal{F}_k^Z\right) = \overline{P}(Y_k = f_j) = \frac{1}{M}. \tag{25}$$

Write

$$\lambda_k = M \sum_{j=1}^{M} \sum_{i=1}^{L} \left\langle C^i \langle Z_k, h_i \rangle X_k, f_j \right\rangle \langle Y_k, f_j \rangle \tag{26}$$

and define

$$\Lambda_k = \prod_{t=0}^{k} \lambda_t. \tag{27}$$

For notational brevity, write $\mathcal{G}_k := \mathcal{F}_k^Y \cup \mathcal{F}_k^X \cup \mathcal{F}_k^Z$. For λ_k, we have the following properties, which help prove the main result on the probability measure change.

Lemma 1: Under probability measure $\overline{P}$ we have

$$\overline{E}\left(\lambda_k | \mathcal{F}_{k-1}^Y \cup \mathcal{F}_k^X \cup \mathcal{F}_k^Z\right) = 1 \tag{28}$$
$$\overline{E}(\lambda_{k+1} | \mathcal{G}_k) = 1. \tag{29}$$

Proof: See Appendix A. ■

The next result indicates the relationship between the estimation problem under P and an unnormalized estimation problem under $\overline{P}$.

Lemma 2 (Theorem 3.2 in [38, ch. 2]): Suppose $(\Omega, \mathcal{F}, P)$ is a probability space and $\mathcal{G} \subset \mathcal{F}$ is a sub- σ-field. Suppose $\hat{P}$ is another probability measure absolutely continuous with respect to P and with Radon–Nikodym derivative $dP/d\hat{P} = \hat{\Lambda}$. If ϕ is any P integrable random variable, then

$$E(\phi|\mathcal{G}) = \begin{cases} \frac{\hat{E}(\hat{\Lambda}\phi|\mathcal{G})}{\hat{E}(\hat{\Lambda}|\mathcal{G})}, & \text{if } \hat{E}(\hat{\Lambda}|\mathcal{G}) > 0 \\ 0, & \text{otherwise.} \end{cases}$$

Based on the above technical lemmas, we are ready to present the main result in this subsection, the proof of which is provided in the Appendix.

Theorem 1: The relationship (15)–(17) under the original probability measure P can be recovered by considering a Radon–Nikodym derivative given by

$$\left.\frac{dP}{d\overline{P}}\right|_{\mathcal{G}_k} = \Lambda_k.$$

Proof: See Appendix B. ■

Remark 1: This result establishes a link between the new probability measure $\overline{P}$ and the original probability measure P. Based on this result, the original joint estimation problem under probability measure P can be solved by solving an unnormalized estimation problem under $\overline{P}$ and mapping the obtained results back to P. Note that Λ_k is related with $X_{1:k}$, $Y_{1:k}$, and $Z_{1:k}$, which are measurable under $\mathcal{G}_k$.

B. Recursive Estimation

Utilizing the change of probability measure approach and the constructed new probability measure $\overline{P}$ in Section III-A, we evaluate the conditional joint distribution of Z_k and X_k under P in this subsection. For notational brevity, we use the Kronecker product and write the unnormalized joint conditional distribution of Z_k and X_k under $\overline{P}$ as

$$q_k := \overline{E}\left[\Lambda_k Z_k \otimes X_k | \mathcal{F}_k^Y\right]. \tag{30}$$

Further, we write the normalized joint conditional distribution under P as

$$p_k := E\left[Z_k \otimes X_k | \mathcal{F}_k^Y\right]. \tag{31}$$

Note that since

$$\sum_{i=1}^{N}\sum_{j=1}^{L}\langle Z_k \otimes X_k, h_i \otimes e_j\rangle = 1$$

and the elements of q_k are nonnegative by the definitions of Λ_k, Z_k and X_k, we have

$$\begin{aligned}\|q_k\|_1 &= \sum_{i=1}^{N}\sum_{j=1}^{L}\langle q_k, h_i \otimes e_j\rangle \\ &= \overline{E}\left[\Lambda_k \sum_{i=1}^{N}\sum_{j=1}^{L}\langle Z_k \otimes X_k, h_i \otimes e_j\rangle \middle| \mathcal{F}_k^Y\right] \\ &= \overline{E}\left[\Lambda_k | \mathcal{F}_k^Y\right].\end{aligned}$$

From Lemma 2, we further have

$$p_k = \frac{q_k}{\|q_k\|_1}. \tag{32}$$

Furthermore, define

$$\alpha_{i,j} := h_i \otimes e_j. \tag{33}$$

We are now in a position to present our main result on the recursive evolution of q_k, the proof of which is deferred to the Appendix.

Theorem 2: The unnormalized joint conditional distribution of Z_k and X_k has the following recursive dynamics under $\overline{P}$:

$$\begin{aligned}q_{k+1} = \operatorname{diag}\left(MY_{k+1}^\top C\right) \cdot (I_L \otimes A) \\ \cdot \left\{\left[\operatorname{diag}\left(\operatorname{vec}(B^\top)\right) \cdot (\mathbf{1}_L \otimes I_L)\right] \otimes I_N\right\} q_k.\end{aligned}$$

Proof: See Appendix C. ■

Remark 2: The above result indicates that, under the new probability measure $\overline{P}$, the unnormalized joint conditional distribution of Z_k and X_k has a recursive form, i.e., q_{k+1} depends solely on q_k rather than $q_{1:k}$. This is encouraging as the computational burden is greatly reduced. On the other hand, the above recursion can be equivalently written as

$$\begin{aligned}&\text{step 1}: q_{k|k} = \left\{\left[\operatorname{diag}\left(\operatorname{vec}(B^\top)\right) \cdot (\mathbf{1}_L \otimes I_L)\right] \otimes I_N\right\} q_k \\ &\text{step 2}: q_{k+1|k} = (I_L \otimes A) q_{k|k} \\ &\text{step 3}: q_{k+1} = \operatorname{diag}\left(MY_{k+1}^\top C\right) q_{k+1|k}.\end{aligned}$$

Despite the complicated hidden relationship among the measurement process $\{Y_k\}$, the hidden process $\{X_k\}$, and the attack process $\{Z_k\}$, the above three-step evolution recursion has clear intuition. The first step corresponds to the attack update, which reflects the effect of the policies of the attacker on the joint distribution. The second step corresponds to a prediction step (which is similar to the prediction step of the Kalman filter [39]), and takes into account of the effect of the state transition of the hidden process $\{X_k\}$. The third step is a measurement update step, and reveals how the measurement information contained in Y_{k+1} is incorporated to update the joint conditional distribution of X_{k+1} and Z_{k+1}, taking advantage of the special structure of the state space of $\{Y_k\}$.

Based on the above recursive expression of q_k, p_k can be obtained according to (32). Let

$$p_k^X := E\left[X_k|\mathcal{F}_k^Y\right] \tag{34}$$

$$p_k^Z := E\left[Z_k|\mathcal{F}_k^Y\right] \tag{35}$$

which denote the conditional probability distributions of X_k and Z_k on $\mathcal{F}_k^Y$, respectively. Based on p_k, p_k^X, and p_k^Z can be directly evaluated as

$$p_k^X = \left(\mathbf{1}_L^\top \otimes I_N\right) p_k,\ p_k^Z = \left(I_L \otimes \mathbf{1}_N^\top\right) p_k. \tag{36}$$

Based on these conditional distributions, the estimates (e.g., the minimum mean square error (MMSE) estimate and the maximum *a posteriori* estimate [40]) can be calculated following standard procedures.

IV. Application to Linear Gaussian Systems

In this section, we apply the proposed results to secure state estimation of linear Gaussian systems. We start with a system with a stable system matrix and present a numerical example of scalar systems. After that, we discuss cases of the vector-valued stable systems as well as the more general nonstationary systems.

A. Example 2 Continued

Firstly, we continue our discussions in Example 2 and introduce a more detailed parameterization of the processes. Consider the following stable linear Gaussian system:

$$x_{k+1} = ax_k + w_k. \tag{37}$$

Here a is the system matrix, x_k is the state with initial state x_0 which is zero mean Gaussian with covariance P_0, and w_k is Gaussian white noise with covariance Q_w. The sensor measurement equation is parameterized as

$$y_k = cx_k + v_k + gz_k \tag{38}$$

where v_k is Gaussian white noise with covariance R and $\{z_k\}$ is a finite-state Markov process with compatible dimensions taking values in $\{\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_L\}$. Define

$$\mathbf{z} := [\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_L]^\top.$$

We assume that x_0, w_k, and v_k are mutually independent. For illustration purposes, we assume that the processes $\{x_k\}$, $\{w_k\}$, and $\{v_k\}$ are scalar-valued. The vector-valued case is discussed later in this section in detail as well. At steady state, $\{x_k\}$ becomes a stationary process with distribution

$$\mu_x(x) = \frac{1}{\sqrt{2\pi Q_x}} \exp\left(-\frac{x^2}{2Q_x}\right) \tag{39}$$

where $Q_x = Q_w/(1-a^2)$. For the scalar case, the quantization regions $\mathbf{x}_i$ and $\mathbf{y}_i$ can be parameterized according to a simple algorithm (see Algorithm 1) as

$$\mathbf{x}_i := \{x \in \mathbb{R} | \underline{x}_i \le x \le \overline{x}_i\},\ \mathbf{y}_i := \{y \in \mathbb{R} | \underline{y}_i \le y \le \overline{y}_i\}.$$

Algorithm 1 Quantizing the support of a scalar signal s into N quantization regions $\mathbf{s}_i$, $i \in \mathbb{N}_{1:N}$

1: Specify the upper and lower limits $\overline{s}$ and $\underline{s}$;
2: $\mathbf{s}_1 := [-\infty, \underline{s}]$;
3: $\mathbf{s}_N := [\overline{s}, \infty]$;
4: Let $s_1 := \underline{s}$, $s_N := \overline{s}$;
5: **for** $i = 2 : N-1$ **do**
6: $\quad s_i := s_{i-1} + (\overline{s} - \underline{s})/(N-2)$;
7: $\quad \mathbf{s}_i := [s_{i-1}, s_i]$;
8: **end for**
9: **end**

Equation (20) becomes

$$a_{i,j}^0 = \int_{\mathbf{x}_j} \int_{\varrho \in \mathbf{x}_i - ax} \mu_w(\varrho) d\varrho \frac{\mu_x(x)}{\int_{\mathbf{x}_j} \mu_x(\xi) d\xi} dx \tag{40}$$

$$= \frac{1}{\int_{\mathbf{x}_j} \mu_x(\xi) d\xi} \int_{\mathbf{x}_j} \left[Q\left(\frac{\underline{x}_i - ax}{\sqrt{Q_w}}\right) - Q\left(\frac{\bar{x}_i - ax}{\sqrt{Q_w}}\right)\right] \times \mu_x(x) dx. \tag{41}$$

Similarly, (22) becomes

$$c_{i,j}^l = \int_{\mathbf{x}_j} \int_{\varrho \in \mathbf{y}_i - cx - \mathbf{z}_l} \mu_v(\varrho) d\varrho \frac{\mu_x(x)}{\int_{\mathbf{x}_j} \mu_x(\xi) d\xi} dx. \tag{42}$$

$$= \frac{1}{\int_{\mathbf{x}_j} \mu_x(\xi) d\xi} \int_{\mathbf{x}_j} \left[Q\left(\frac{\underline{y}_i - cx - \mathbf{z}_l}{\sqrt{Q_v}}\right) - Q\left(\frac{\bar{y}_i - cx - \mathbf{z}_l}{\sqrt{Q_v}}\right)\right] \mu_x(x) dx. \tag{43}$$

As $\mathbf{x}_j$, $\mathbf{y}_i$, and μ_x are time invariant, the values of $a_{i,j}^0$ and $c_{i,j}^0$ can be determined offline using standard numerical integration techniques. The outcome is analyzed and compared in the numerical example.

For numerical illustration, take $a = 0.9$, $c = 0.5$, $g = 1$, $Q_w = 1$, and $R = 1$. Following the fashion of standard analog-to-digital converters (ADCs) as in Algorithm 1, we quantize the support of x_k into N quantization regions choosing $\underline{x} = -6$ and $\bar{x} = 6$, and the support of y_k into M quantization regions choosing $\underline{y} = -5$ and $\bar{y} = 5$. Based on the obtained HMM model, the results in Theorem 2 are applied, and the HMM-based MMSE estimates, which we refer to as "HMM-based estimates" hereafter, are obtained as

$$\hat{x}_k = \check{\mathbf{x}}^\top p_k^X$$

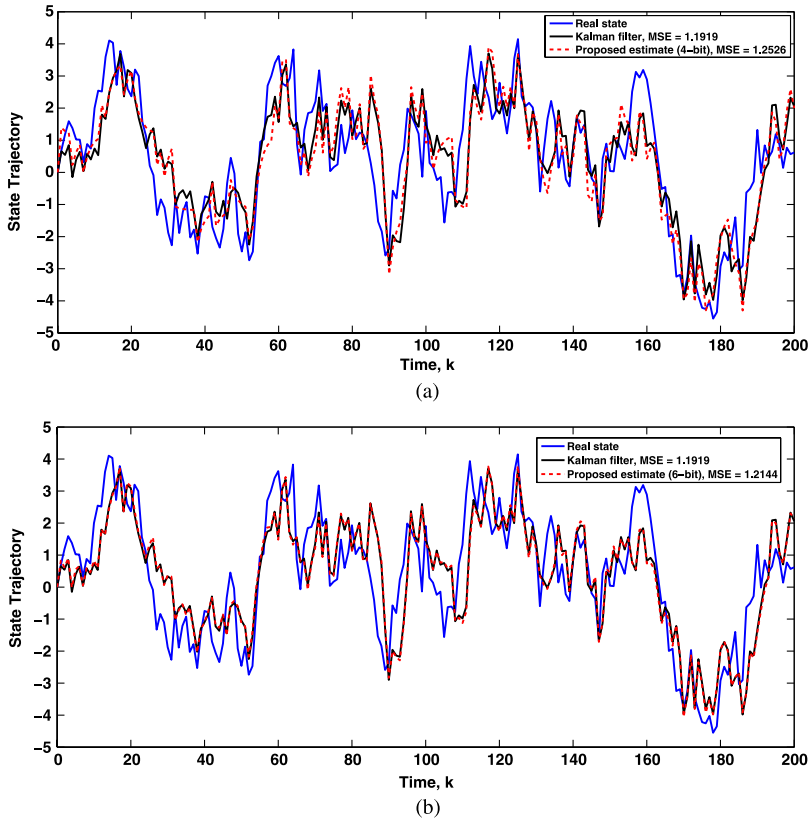


Fig. 4. Comparison with Kalman filter. (a) The 4-b estimates. (b) The 6-b estimates.

where $\check{\mathbf{x}} := [\check{x}_1, \ldots, \check{x}_N]^\top$ with $\check{x}_1 = \bar{x}$, $\check{x}_N = \underline{x}$ and $\check{x}_i = (\bar{x}_i + \underline{x}_i)/2$. Similarly, the estimate for z_k is obtained as

$$\hat{z}_k = \mathbf{z}^\top p_k^Z.$$

Before looking into the attacked case, we first evaluate the consequence of quantization by taking $\mathbf{z}_i = 0$ for all $i \in \mathbb{R}_{1:L}$ and comparing the HMM-based estimates with the Kalman filter, namely, the exact MMSE estimates with no quantization effect. As $\mathbf{z}_i = 0$ for all i, the parameters in B do not affect the results. To analyze the effect of quantization, we consider two choices of N and M, i.e., 1) $N = M = 16$ and 2) $N = M = 64$; these choices correspond to a 4-b ADC and a 6-b ADC, respectively.[4] The results are shown in Fig. 4. Even for the 4-b case, the HMM-based estimate stays close to the Kalman filter with small perceptible differences, and effectively approximates the unknown real state. When the number of bits increases to 6, which still corresponds to a low-resolution ADC, the difference between the HMM-based MMSE estimate and the Kalman filter is hardly distinguishable. To quantitatively measure the differences, the actual mean square error (MSE) values are also provided in the legends.

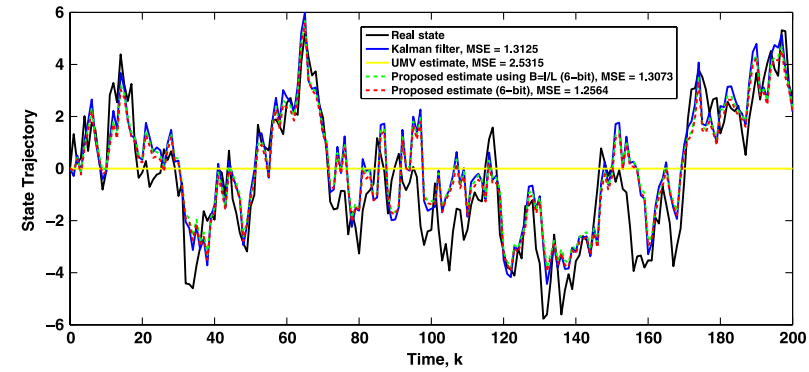
Now we consider the case with attacks. Apart from the Kalman filter, two more estimators are incorporated for comparison in this case as well. The first estimator is also an HMM-based estimator calculated according to the result in Theorem 2, but this estimator is calculated by assuming B is not known, so that a noninformative guess B_0 of B satisfying

$$B_0 = \frac{\mathbf{1}_L}{L} \tag{44}$$

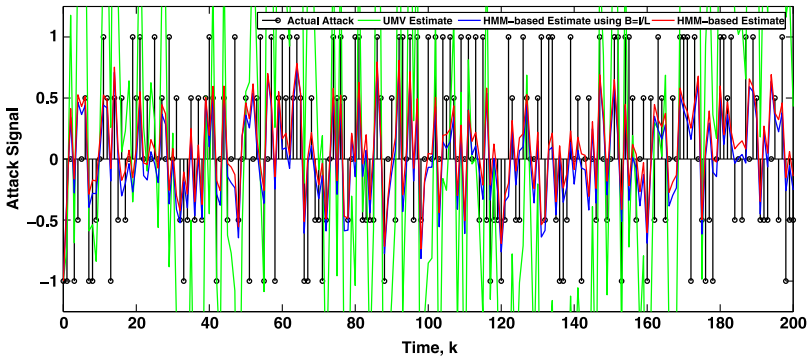
is utilized to calculate the estimates for the state and attack processes, where $\mathbf{1}_L \in \mathbb{R}^{L\times L}$ denotes a matrix with all its elements equal to 1. Note that the choice $B_0 = \mathbf{1}_L/L$ implies that the estimator has no knowledge about the decision process of the attacker, which is the case when the transition probability matrix B is not known. The second estimator considered is the unbiased minimum variance (UMV) estimator introduced in [42] that provides optimal estimates $\tilde{x}_k$ and $\tilde{z}_k$ for the states x_k and unknown exogenous inputs z_k in the sense of minimum variance and unbiasedness for discrete-time linear Gaussian systems in the form of equations (37) and (38). Specifically, the estimator equations are given by

$$\tilde{x}_{k+1} = a\tilde{x}_k + K_k(y_k - ca\tilde{x}_k - g\tilde{z}_k) \tag{45}$$
$$\tilde{z}_k = M_k(y_k - ca\tilde{x}_k) \tag{46}$$
$$K_k = P_k^- c^\top \tilde{R}_k^{-1} \tag{47}$$
$$M_k = \left(g^\top \tilde{R}_k^{-1} g\right)^{-1} g^\top \tilde{R}_k^{-1} \tag{48}$$
$$\tilde{R}_k = cP_k^- c^\top + R \tag{49}$$
$$P_k^- = aP_k a^\top + Q_w \tag{50}$$
$$P_k = P_k^- - K_k\left[\tilde{R}_k - g(g^\top \tilde{R}_k^{-1} g)^{-1}\right] g^\top K_k^\top. \tag{51}$$

[4]Note that the bit settings here can be fulfilled by most prevailing ADCs, which are normally 8–16 b [41].



(a)



(b)

Fig. 5. Results for the balanced attack case. (a) Estimation results for the state trajectory. (b) Estimation results for the attack signal.

This estimator was developed by ignoring the potential decision dynamics of the attack process z_k and treating it as an unknown deterministic input term, and the computation complexity is similar to that of the Kalman filter. This type of estimator (which we will name as "UMV estimator" hereafter) has been recently utilized for input reconstruction for networked systems subject to deception attacks [43].

First, we assume $L = 5$ and

$$\mathbf{z} = [-1,\ -0.5,\ 0,\ 0.5,\ 1]^\top. \tag{52}$$

The transition matrix B is randomly generated as

$$B = \begin{bmatrix} 0.1503 & 0.1807 & 0.0522 & 0.2741 & 0.0967 \\ 0.3347 & 0.1714 & 0.0893 & 0.2802 & 0.1737 \\ 0.2835 & 0.3289 & 0.0468 & 0.2011 & 0.3389 \\ 0.1652 & 0.0390 & 0.3765 & 0.0166 & 0.2277 \\ 0.0663 & 0.2800 & 0.4353 & 0.2279 & 0.1631 \end{bmatrix}.$$

The results are plotted in Fig. 5 with the MSE values included. The Kalman filter, the HMM-based estimator using the actual B matrix, and the HMM-based estimator using $B = B_0$ achieve similar performance in the sense of MSE, although the MSE value of the HMM-based estimator using the actual B matrix is a little bit better than those of the other two estimators. The explanation is that when the support of the attack signal is balanced [see (52)], the signal approximately behaves like a zero mean noise, to which Kalman filter has a good robust property [44]. This property no longer holds when the support of the attack signal becomes unbalanced. To see this, we let

$$\mathbf{z} = [-1,\ -0.5,\ 0,\ 0.5,\ 1,\ 2,\ 3]^\top \tag{53}$$

and correspondingly a transition matrix is generated randomly as

$$B = \begin{bmatrix} 0.226 & 0.014 & 0.040 & 0.013 & 0.230 & 0.030 & 0.166 \\ 0.311 & 0.095 & 0.181 & 0.235 & 0.078 & 0.045 & 0.106 \\ 0.100 & 0.375 & 0.156 & 0.075 & 0.360 & 0.213 & 0.083 \\ 0.014 & 0.315 & 0.029 & 0.014 & 0.035 & 0.137 & 0.080 \\ 0.099 & 0.065 & 0.238 & 0.340 & 0.000 & 0.075 & 0.048 \\ 0.013 & 0.117 & 0.168 & 0.044 & 0.262 & 0.035 & 0.339 \\ 0.236 & 0.020 & 0.188 & 0.279 & 0.036 & 0.466 & 0.179 \end{bmatrix}.$$

The results are plotted in Fig. 6 with the MSE values included. In this case, the estimates obtained by the Kalman filter are apparently biased and, specifically, are larger than the true values of the states, resulting in an MSE value of 2.5684, due to the unconsidered attack signal with a positive-oriented support. The HMM-based estimates (for both the case of using the actual B matrix and using $B = B_0$), on the other hand, stay relatively closer to the true states, with MSE values being 1.6588 and 1.7368. Also, notice that for both attack scenarios, the attack signals can be efficiently estimated by the two HMM-based estimators, especially when the cardinality of the state space of the attack process is relatively large.

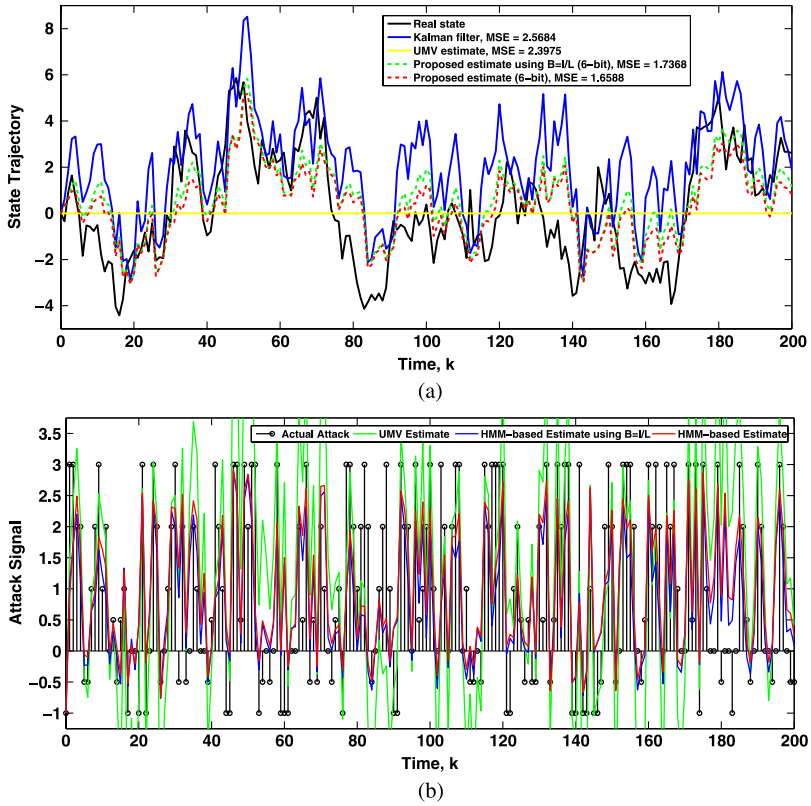


Fig. 6. Results for the unbalanced attack case. (a) Estimation results for the state trajectory. (b) Estimation results for the attack signal.

For both choices of the attack processes z_k, we observe the state estimates of the UMV estimator are always 0 (although the values of the estimated attack signal are nonzero). The reason is that when g is nonsingular, the innovation term $y_k - ca\tilde{x}_k - g\tilde{z}_k$ in (45) becomes

$$\begin{aligned} y_k - ca\tilde{x}_k - g\tilde{z}_k &= y_k - ca\tilde{x}_k - gM_k(y_k - ca\tilde{x}_k) \\ &= y_k - ca\tilde{x}_k - g\left(g^\top \tilde{R}_k^{-1} g\right)^{-1} \\ &\quad \times g^\top \tilde{R}_k^{-1}(y_k - ca\tilde{x}_k) \\ &= 0. \end{aligned}$$

In this way, since a is stable, the solution to equation (45) converges to 0 exponentially, and leads to comparatively large estimation errors; as a consequence, the estimates for the attack signal become relatively inaccurate as well. Another note to make is that the proposed HMM estimator possesses a resilience property in the sense that the estimation performance is robust with respect to the choice of B matrix; from Figs. 5 and 6, it is observed the estimates for both the state signals and the attack signals generated by the HMM estimator using an noninformative choice of B (namely, $B = B_0 = I_L/L$) are close to those of the HMM estimator obtained using the actual B matrix. This property is helpful in secure estimation, as normally the B matrix cannot be accurately known or estimated.

B. Further Discussions

In this section, a few necessary discussions are presented on applying the proposed estimator to general linear Gaussian systems. First, we focus on the unattacked case and make a comparison between the Kalman filter and the HMM-based estimator. In engineering systems, quantization errors always exist in sensor measurements, so that the true values of the measurements cannot be known and only the quantized values are available. Using the notation introduced in Section IV-A, the actual available quantized measurement information up to time k is recursively defined by $\mathcal{I}_0 := \{y_0 \in \mathbf{y}_{i_0}\}$

$$\mathcal{I}_k := \{y_k \in \mathbf{y}_{i_k}\} \cup \mathcal{I}_{k-1}$$

where i_k denotes the index of the quantization region in which y_k belongs to at time instant k. Let

$$\mathcal{Y}_k := \{y_0, y_1, \ldots, y_k\}$$

denote the set of true values of the measurements up to time k, and let

$$\hat{\mathcal{Y}}_k := \{\hat{y}_k, \hat{y}_1, \ldots, \hat{y}_k\}$$

denote the set of quantized values $\hat{y}_k$ of the measurements up to time k. Based on the notation introduced, the problem of optimal state estimation is to find the MMSE estimate $\bar{x}_k$ of x_k based on the available measurement information $\mathcal{I}_k$. From the standard results in signal estimation theory [39], [40], this MMSE estimate is actually the mean of the conditional distribution of x_k on $\mathcal{I}_k$, namely

$$\bar{x}_k = E(x_k|\mathcal{I}_k) \tag{54}$$

$$= \sum_{i=1}^{N} \int_{x \in \mathbf{x}_i} x f_{x_k}(x|\mathcal{I}_k) dx \tag{55}$$

which normally cannot be expressed by a simple expression. The Kalman filter was originally developed on the basis of the set of true measurement values $\mathcal{Y}_k$ and calculates $\bar{\bar{x}}_k = E(x_k|\mathcal{Y}_k)$ recursively, which is the MMSE estimate given $\mathcal{Y}_k$, but it cannot handle set-valued measurement information of the form $\{y_k \in \mathbf{y}_{i_k}\}$. When a Kalman filter is applied, the estimation problem based on quantized measurement information is solved in an indirect way, and the estimates (which are actually $E(x_k|\hat{\mathcal{Y}}_k)$) are calculated based on the set of quantized measurements $\hat{\mathcal{Y}}_k$ instead of $\mathcal{Y}_k$. The relationship between $E(x_k|\hat{\mathcal{Y}}_k)$ and $E(x_k|\mathcal{I}_k)$, however, is difficult to analyze in general. On the other hand, the HMM-based estimator evaluates the conditional distribution $P(x_k \in \mathbf{x}_i|\mathcal{I}_k)$, and the HMM-based estimates are calculated as

$$\begin{aligned}\hat{x}_k &= \check{\mathbf{x}}^\top p_k^X \\ &= \sum_{i=1}^{N} \check{x}_i P(x_k \in \mathbf{x}_i|\mathcal{I}_k) \qquad (56)\\ &= \sum_{i=1}^{N} \int_{x\in\mathbf{x}_i} \check{x}_i f_{x_k}(x|\mathcal{I}_k)dx \qquad (57)\end{aligned}$$

where we recall that $\check{x}_i$ denotes the quantized value for $x \in \mathbf{x}_i$ such that the output of the quantizer equals $\check{x}_i$ for all $x \in \mathbf{x}_i$. A comparison between (55) and (57) indicates that if $\check{x}_i$ are properly chosen, we have $\hat{x}_k \to \bar{x}_k$ as $N \to \infty$, $\bar{x}_k$ being the MMSE estimate of x_k based on the quantized measurement information $\mathcal{I}_k$.

Intuitively, the modeling procedure introduced in Section IV-A applies almost equally to vector-valued stationary Gaussian processes, as the quantization procedure in Algorithm 1 can be easily extended to the case of vector-valued systems; but the difference is that (41) and (43) no longer hold, and instead we only have (40) and (42) to calculate $a^0_{i,j}$ and $c^l_{i,j}$, for which the integrations with higher dimensions need to be evaluated through numerical integration. Based on the values of $a^0_{i,j}$ and $c^l_{i,j}$, the results developed for secure estimation (namely, Theorem 2) can be directly applied, so that the HMM-based estimates $\hat{x}_k$ can be obtained. The main differences in considering vector-valued stationary Gaussian processes include: 1) the numerical integration algorithm to calculate $a^0_{i,j}$ and $c^l_{i,j}$ needs to be smartly designed as multiple-dimensional numerical integration problems need to be solved; and 2) the corresponding computational complexity is increased. Fortunately, this increased computational burden can be mostly absorbed offline, due to the stationarity and constant parameter matrices. As the implementation of the results in Theorem 2 only requires simple matrix computations, the increase of the online computation burden is that matrix computations with higher dimensions will be needed. The increase of the online computation burden, however, is manageable in general, because: 1) a relatively small number of quantization regions for each state would be necessary for satisfactory estimation performance, as has been shown in the numerical example; and 2) the computation burden of matrix calculations is much smaller compared with that of evaluating the multiple-dimensional numerical integrations, and normally stays within the capability of digital signal processors, which is increasing at a dramatic speed nowadays due to the developments in microelectronics.

In principle, the results proposed can be applied to non-stationary processes as well, but the cost is that the assumed offline calculation of $a^0_{i,j}$ and $c^l_{i,j}$ has to be done online, as these parameters become time-varying. Some interesting and important exceptions include stable linear periodic Gaussian systems, for which the corresponding $a^0_{i,j}$ and $c^l_{i,j}$ parameters also change in time but have only a finite number of possible parameterizations, which can still be calculated offline. For practitioners, the state estimation of unstable Gaussian systems with quantized measurements, e.g., using ADCs such as AD7988 [41] and digital microprocessors, is a problem that goes beyond estimator design itself—the states inevitably blow up and exceed all bounded limits of the quantizers in some finite time. Thus figuring out the appropriate quantization algorithm already becomes an extraordinarily challenging problem, which hampers the subsequent estimator design and implementation procedure.

V. Conclusion

In this paper, a stochastic modeling framework has been proposed for CPSs considering the effect of adversary attacks. Based on this framework, a joint state and attack estimation problem has been formulated and solved, utilizing the change of probability measure approach. The results have been applied to stable linear Gaussian systems based on the quantization effects of ADC, and extensions to more general systems have also been discussed. Based on the obtained results, two immediate important problems to be investigated include the parameter estimation problem for the attack process and the joint state and parameter estimation problem. In the current work, the secure estimation problem is approached from the estimator side; an alternative way of considering secure estimation is to exploit the fact that the estimator and the attacker may be aware of each other so that the decision-making process will become interactive and depend on both parties. In this scenario, game-theoretic approaches [8], [45] need to be utilized to solve the underlying joint state and attack estimation problems, in which case the transition probability matrix of the attacker as well as the state estimates will be determined to guarantee the worst-case performance. These extensions provide the directions for our future work.

Appendix

A. Proof of Lemma 1

By definition

$$\begin{aligned}&\overline{E}\left(\lambda_k|\mathcal{F}^Y_{k-1}\cup\mathcal{F}^X_k\cup\mathcal{F}^Z_k\right)\\ &\quad=\overline{E}\left(M\sum_{j=1}^{M}\sum_{i=1}^{L}\left\langle C^i\langle Z_k,h_i\rangle X_k,f_j\right\rangle\langle Y_k,f_j\rangle\right.\\ &\qquad\left.|\mathcal{F}^Y_{k-1}\cup\mathcal{F}^X_k\cup\mathcal{F}^Z_k\right).\end{aligned}$$

Since Y_k is independent of $Y_{1:k-1}$, X_k and Z_k under $\overline{P}$, we have

$$\begin{aligned}&\overline{E}\left(\lambda_k|\mathcal{F}_{k-1}^Y\cup\mathcal{F}_k^X\cup\mathcal{F}_k^Z\right)\\&\quad=M\sum_{j=1}^M\sum_{i=1}^L\overline{E}\left(\langle Y_k,f_j\rangle\right)\overline{E}\\&\qquad\left(\left\langle C^i\langle Z_k,h_i\rangle X_k,f_j\right\rangle|\mathcal{F}_{k-1}^Y\cup\mathcal{F}_k^X\cup\mathcal{F}_k^Z\right).\end{aligned}$$

Also, since Y_k follows a uniform distribution we have:

$$\overline{E}\left(\langle Y_k,f_j\rangle\right)=\frac{1}{M}.$$

As X_k and Z_k are measurable under F_k^X and F_k^Z, it follows that:

$$\begin{aligned}&\overline{E}\left(\langle C^i\langle Z_k,h_i\rangle X_k,f_j\rangle|\mathcal{F}_{k-1}^Y\cup\mathcal{F}_k^X\cup\mathcal{F}_k^Z\right)\\&\qquad=\left\langle\langle Z_k,h_i\rangle\sum_{j_0=1}^N c_{j_0}^i\langle X_k,e_{j_0}\rangle,f_j\right\rangle.\end{aligned}$$

Therefore we have

$$\begin{aligned}&\overline{E}\left(\lambda_k|\mathcal{F}_{k-1}^Y\cup\mathcal{F}_k^X\cup\mathcal{F}_k^Z\right)\\&\quad=\sum_{j=1}^M\sum_{i=1}^L\left\langle\langle Z_k,h_i\rangle\sum_{j_0=1}^N c_{j_0}^i\langle X_k,e_{j_0}\rangle,f_j\right\rangle\\&\quad=\sum_{j=1}^M\sum_{i=1}^L\langle Z_k,h_i\rangle\sum_{j_0=1}^N c_{j,j_0}^i\langle X_k,e_{j_0}\rangle\\&\quad=\sum_{i=1}^L\langle Z_k,h_i\rangle\sum_{j_0=1}^N\langle X_k,e_{j_0}\rangle=1.\end{aligned}$$

For $\overline{E}(\lambda_{k+1}|\mathcal{G}_k)$, we have

$$\begin{aligned}&\overline{E}(\lambda_{k+1}|\mathcal{G}_k)\\&=\overline{E}\left(M\sum_{j=1}^M\sum_{i=1}^L\left\langle C^i\langle Z_{k+1},h_i\rangle X_{k+1},f_j\right\rangle\langle Y_{k+1},f_j\rangle\middle|\mathcal{G}_k\right)\\&=\overline{E}\left(\sum_{j=1}^M\sum_{i=1}^L\left\langle C^i\langle Z_{k+1},h_i\rangle X_{k+1},f_j\right\rangle\middle|\mathcal{G}_k\right)\end{aligned}$$

as Y_k is i.i.d. uniformly distributed under $\overline{P}$. Since X_k takes value in S_X, which is composed of the unit vectors in $\mathbb{R}^N$

$$C^iX_{k+1}=\sum_{i_0=1}^N\langle X_{k+1},e_{i_0}\rangle C^ie_{i_0}=\sum_{i_0=1}^N\langle X_{k+1},e_{i_0}\rangle c_{i_0}^i.$$

In this way, following the fact that $c_{j,i_0}^i=\langle c_{i_0}^i,f_j\rangle$ and $\sum_{j=1}^M c_{j,i_0}^i=1$, we have:

$$\begin{aligned}&\overline{E}(\lambda_{k+1}|\mathcal{G}_k)\\&\quad=\overline{E}\left(\sum_{j=1}^M\sum_{i=1}^L\sum_{i_0=1}^N c_{j,i_0}^i\langle X_{k+1},e_{i_0}\rangle\langle Z_{k+1},h_i\rangle\middle|\mathcal{G}_k\right)\\&\quad=\sum_{i=1}^L\sum_{i_0=1}^N\overline{E}\left(\sum_{j=1}^M c_{j,i_0}^i\langle X_{k+1},e_{i_0}\rangle\langle Z_{k+1},h_i\rangle\middle|\mathcal{G}_k\right)\\&\quad=\sum_{i=1}^L\sum_{i_0=1}^N\overline{E}\left(\langle X_{k+1},e_{i_0}\rangle\langle Z_{k+1},h_i\rangle|\,\mathcal{G}_k\right)=1\end{aligned}$$

which completes the proof.

B. Proof of Theorem 1

Since $\lambda_{1:k}$ are measurable in $\mathcal{G}_k$, we have

$$\overline{E}(\Lambda_{k+1}X_{k+1}|\mathcal{G}_k)=\left(\prod_{t=1}^k\lambda_t\right)\overline{E}(\lambda_{k+1}X_{k+1}|\mathcal{G}_k)$$

and $\overline{E}(\Lambda_{k+1}|\mathcal{G}_k)=(\prod_{t=1}^k\lambda_t)\overline{E}(\lambda_{k+1}|\mathcal{G}_k)=\prod_{t=1}^k\lambda_t$, as $\overline{E}(\lambda_{k+1}|\mathcal{G}_k)=1$ according to Lemma 1. Lemma 2 and the definition of λ_{k+1} imply

$$\begin{aligned}E(X_{k+1}|\mathcal{G}_k)&=\frac{\overline{E}(\Lambda_{k+1}X_{k+1}|\mathcal{G}_k)}{\overline{E}(\Lambda_{k+1}|\mathcal{G}_k)}\\&=\overline{E}(\lambda_{k+1}X_{k+1}|\mathcal{G}_k)\\&=\overline{E}\left(M\sum_{j=1}^M\sum_{i=1}^L\left\langle C^i\langle Z_{k+1},h_i\rangle X_{k+1},f_j\right\rangle\right.\\&\qquad\left.\times\,\langle Y_{k+1},f_j\rangle X_{k+1}|\mathcal{G}_k\right).\end{aligned}$$

Following a similar argument as in the proof of Lemma 1, we further have:

$$\begin{aligned}&E(X_{k+1}|\mathcal{G}_k)\\&=\overline{E}\left(\sum_{j=1}^M\sum_{i=1}^L\left\langle\sum_{i_0=1}^N c_{i_0}^i\langle X_{k+1},e_{i_0}\rangle\langle Z_{k+1},h_i\rangle,f_j\right\rangle X_{k+1}|\,\mathcal{G}_k\right)\\&=\overline{E}\left[\sum_{i_0=1}^N\sum_{i=1}^L\left(\sum_{j=1}^M c_{j,i_0}^i\langle X_{k+1},e_{i_0}\rangle\langle Z_{k+1},h_i\rangle\right)X_{k+1}|\mathcal{G}_k\right]\\&=\overline{E}\left[\sum_{i_0=1}^N\sum_{i=1}^L\langle X_{k+1},e_{i_0}\rangle\langle Z_{k+1},h_i\rangle X_{k+1}\middle|\mathcal{G}_k\right]\\&=\overline{E}\left[\sum_{i=1}^L\langle Z_{k+1},h_i\rangle X_{k+1}\middle|\mathcal{G}_k\right].\end{aligned}$$

By repeated conditioning, we have

$$\begin{aligned}
&\overline{E}\left[\sum_{i=1}^{L}\langle Z_{k+1},h_i\rangle X_{k+1}\middle|\mathcal{G}_k\right]\\
&\quad=\overline{E}\left[\overline{E}\left(\sum_{i=1}^{L}\langle Z_{k+1},h_i\rangle X_{k+1}\middle|\mathcal{G}_k\cup X_{k+1}\right)\middle|\mathcal{G}_k\right]\\
&\quad=\overline{E}\left[X_{k+1}\overline{E}\left(\sum_{i=1}^{L}\langle Z_{k+1},h_i\rangle\middle|\mathcal{G}_k\cup X_{k+1}\right)\middle|\mathcal{G}_k\right]\\
&\quad=\overline{E}(X_{k+1}|\mathcal{G}_k)=\sum_{i=1}^{L}A^i\langle Z_k,h_i\rangle X_k
\end{aligned}$$

which recovers (17).

Following a similar procedure, we have:

$$\begin{aligned}
&E(Z_{k+1}|\mathcal{G}_k)\\
&=\frac{\overline{E}(\Lambda_{k+1}Z_{k+1}|\mathcal{G}_k)}{\overline{E}(\Lambda_{k+1}|\mathcal{G}_k)}\\
&=\overline{E}(\lambda_{k+1}Z_{k+1}|\mathcal{G}_k)\\
&=\overline{E}\left[\sum_{i_0=1}^{N}\sum_{i=1}^{L}\left(\sum_{j=1}^{M}c^i_{j,i_0}\langle X_{k+1},e_{i_0}\rangle\langle Z_{k+1},h_i\rangle\right)Z_{k+1}\middle|\mathcal{G}_k\right]\\
&=\overline{E}\left[\sum_{i_0=1}^{N}\sum_{i=1}^{L}\langle X_{k+1},e_{i_0}\rangle\langle Z_{k+1},h_i\rangle Z_{k+1}\middle|\mathcal{G}_k\right]\\
&=\overline{E}\left[\sum_{i_0=1}^{N}\langle X_{k+1},e_{i_0}\rangle Z_{k+1}\middle|\mathcal{G}_k\right]\\
&=\overline{E}(Z_{k+1}|\mathcal{G}_k)=BZ_k
\end{aligned}$$

which recovers (15). Also

$$\begin{aligned}
&\overline{E}\left(Y_k|\mathcal{F}^Y_{k-1}\cup\mathcal{F}^X_k\cup\mathcal{F}^Z_k\right)\\
&\quad=\frac{\overline{E}\left(\Lambda_kY_k|\mathcal{F}^Y_{k-1}\cup\mathcal{F}^X_k\cup\mathcal{F}^Z_k\right)}{\overline{E}\left(\Lambda_k|\mathcal{F}^Y_{k-1}\cup\mathcal{F}^X_k\cup\mathcal{F}^Z_k\right)}\\
&\quad=\overline{E}\left(\lambda_kY_k|\mathcal{F}^Y_{k-1}\cup\mathcal{F}^X_k\cup\mathcal{F}^Z_k\right)\\
&\quad=\overline{E}\left(M\sum_{j=1}^{M}\sum_{i=1}^{L}\left\langle C^i\langle Z_k,h_i\rangle X_k,f_j\right\rangle\langle Y_k,f_j\rangle Y_k\right.\\
&\qquad\left.|\mathcal{F}^Y_{k-1}\cup\mathcal{F}^X_k\cup\mathcal{F}^Z_k\right)\\
&\quad=M\sum_{j=1}^{M}\sum_{i=1}^{L}\left\langle C^i\langle Z_k,h_i\rangle X_k,f_j\right\rangle\overline{E}\left(\langle Y_k,f_j\rangle Y_k\right)\\
&\quad=\sum_{j=1}^{M}\sum_{i=1}^{L}\left\langle C^i\langle Z_k,h_i\rangle X_k,f_j\right\rangle f_j\\
&\quad=\sum_{i=1}^{L}C^i\langle Z_k,h_i\rangle X_k
\end{aligned}$$

which recovers (16) and completes the proof.

C. Proof of Theorem 2

The proof of Theorem 2 is much involved. Before continuing, we first prove the following technical lemma.

Lemma 3:

$$\begin{aligned}
&\overline{E}\left[\Lambda_k\left\langle Z_{k+1}\otimes X_{k+1},\alpha_{i_0,j_0}\right\rangle|\mathcal{F}^Y_{k+1}\right]\\
&\qquad=\sum_{i_1=1}^{L}\sum_{j_1=1}^{N}\left\langle B_{i_1}\otimes A^{i_1}_{j_1},\alpha_{i_0,j_0}\right\rangle\left\langle q_k,\alpha_{i_1,j_1}\right\rangle.
\end{aligned}$$

Proof: By repeated conditioning, we have

$$\begin{aligned}
&\overline{E}\left[\Lambda_k\langle Z_{k+1}\otimes X_{k+1},\alpha_{i_0,j_0}\rangle|\,\mathcal{F}^Y_{k+1}\right]\\
&\quad=\overline{E}\left[\overline{E}\left[\Lambda_k\langle Z_{k+1}\otimes X_{k+1},\alpha_{i_0,j_0}\rangle|\,\mathcal{F}^Y_{k+1}\cup\mathcal{F}^X_{k+1}\right.\right.\\
&\qquad\left.\left.\cup\mathcal{F}^Z_k\right]\middle|\,\mathcal{F}^Y_{k+1}\right]\\
&\quad=\overline{E}\left[\Lambda_k\langle\overline{E}\left(Z_{k+1}|\,\mathcal{F}^Y_{k+1}\cup\mathcal{F}^X_{k+1}\cup\mathcal{F}^Z_k\right)\otimes X_{k+1},\right.\\
&\qquad\left.\alpha_{i_0,j_0}\rangle|\,\mathcal{F}^Y_{k+1}\right]
\end{aligned}\tag{58}$$

where (58) is due to the fact that Λ_k is measurable under $\mathcal{F}^Y_{k+1}\cup\mathcal{F}^X_{k+1}\cup\mathcal{F}^Z_k$. Since $\overline{E}(Z_{k+1}|\mathcal{F}^Y_{k+1}\cup\mathcal{F}^X_{k+1}\cup\mathcal{F}^Z_k)=BZ_k$, we further have

$$\begin{aligned}
&\overline{E}\left[\Lambda_k\langle Z_{k+1}\otimes X_{k+1},\alpha_{i_0,j_0}\rangle|\,\mathcal{F}^Y_{k+1}\right]\\
&\qquad=\overline{E}\left[\Lambda_k\langle(BZ_k)\otimes X_{k+1},\alpha_{i_0,j_0}\rangle|\,\mathcal{F}^Y_{k+1}\right].
\end{aligned}\tag{59}$$

Noticing that

$$\begin{aligned}
&\overline{E}\left[\Lambda_k\left\langle(BZ_k)\otimes X_{k+1},\alpha_{i_0,j_0}\right\rangle|\,\mathcal{F}^Y_{k+1}\right]\\
&\quad=\overline{E}\left[\overline{E}\left[\Lambda_k\langle(BZ_k)\otimes X_{k+1},\alpha_{i_0,j_0}\rangle|\,\mathcal{F}^Y_{k+1}\cup\mathcal{F}^Z_k\right.\right.\\
&\qquad\left.\left.\cup\mathcal{F}^X_k\right]\middle|\,\mathcal{F}^Y_{k+1}\right]\\
&\quad=\overline{E}\left[\Lambda_k\left\langle(BZ_k)\otimes\overline{E}\left[X_{k+1}|\,\mathcal{F}^Y_{k+1}\cup\mathcal{F}^Z_k\cup\mathcal{F}^X_k\right],\right.\right.\\
&\qquad\left.\left.\alpha_{i_0,j_0}\right\rangle|\,\mathcal{F}^Y_{k+1}\right]\\
&\quad=\overline{E}\left[\Lambda_k\left\langle(BZ_k)\otimes\left[\sum_{i=1}^{L}A^i\langle Z_k,h_i\rangle X_k\right],\alpha_{i_0,j_0}\right\rangle\middle|\,\mathcal{F}^Y_{k+1}\right]
\end{aligned}$$

we have that (59) reduces to

$$\begin{aligned}
&\overline{E}\left[\Lambda_k\langle Z_{k+1}\otimes X_{k+1},\alpha_{i_0,j_0}\rangle|\,\mathcal{F}^Y_{k+1}\right]\\
&\quad=\overline{E}\left[\Lambda_k\left\langle(BZ_k)\otimes\left[A(Z_k\otimes X_k)\right],\alpha_{i_0,j_0}\right\rangle|\,\mathcal{F}^Y_{k+1}\right].
\end{aligned}\tag{60}$$

Since Z_k takes value from $\{h_i\}$, we have

$$\begin{aligned}
&(BZ_k)\otimes\left[A(Z_k\otimes X_k)\right]\\
&\quad=\sum_{i_1=1}^{L}(Bh_{i_1})\otimes\left[A\sum_{j_1=1}^{N}\langle Z_k\otimes X_k,\alpha_{i_1,j_1}\rangle\alpha_{i_1,j_1}\right]\\
&\quad=\sum_{i_1=1}^{L}\sum_{j_1=1}^{N}B_{i_1}\otimes\left[\langle Z_k\otimes X_k,\ \alpha_{i_1,j_1}\rangle A^{i_1}_{j_1}\right]
\end{aligned}$$

based on which (60) further becomes

$$\begin{aligned}
&\overline{E}\left[\Lambda_k\left\langle Z_{k+1}\otimes X_{k+1},\alpha_{i_0,j_0}\right\rangle|\,\mathcal{F}_{k+1}^{Y}\right]\\
&\quad=\sum_{i_1=1}^{L}\sum_{j_1=1}^{N}\Big\langle B_{i_1}\otimes\Big[\big\langle\overline{E}\left[\Lambda_k Z_k\otimes X_k|\mathcal{F}_{k+1}^{Y}\right],\\
&\qquad\qquad\alpha_{i_1,j_1}\big\rangle\,A_{j_1}^{i_1}\Big],\alpha_{i_0,j_0}\Big\rangle\\
&\quad=\sum_{i_1=1}^{L}\sum_{j_1=1}^{N}\left\langle B_{i_1}\otimes\left\langle q_k,\alpha_{i_1,j_1}\right\rangle A_{j_1}^{i_1},\alpha_{i_0,j_0}\right\rangle\\
&\quad=\sum_{i_1=1}^{L}\sum_{j_1=1}^{N}\left\langle B_{i_1}\otimes A_{j_1}^{i_1},\alpha_{i_0,j_0}\right\rangle\left\langle q_k,\alpha_{i_1,j_1}\right\rangle
\end{aligned}$$

which completes the proof. ■

Based on the above result, we further prove Theorem 2. According to the definition of q_{k+1} and noticing that

$$\sum_{i=1}^{L}C^i\langle Z_{k+1},h_i\rangle X_{k+1}=C(Z_{k+1}\otimes X_{k+1})$$

we have

$$\begin{aligned}
q_{k+1}&=\overline{E}\left[\Lambda_{k+1}Z_{k+1}\otimes X_{k+1}|\mathcal{F}_{k+1}^{Y}\right]\\
&=\overline{E}\left[\Lambda_k M\sum_{j=1}^{M}\sum_{i=1}^{L}\left\langle C^i\langle Z_{k+1},h_i\rangle X_{k+1},f_j\right\rangle\right.\\
&\qquad\left.\times\left\langle Y_{k+1},f_j\right\rangle\;Z_{k+1}\otimes X_{k+1}|\,\mathcal{F}_{k+1}^{Y}\right] \qquad (61)\\
&=\overline{E}\left[\Lambda_k M\sum_{j=1}^{M}\left\langle C(Z_{k+1}\otimes X_{k+1}),f_j\right\rangle\left\langle Y_{k+1},f_j\right\rangle\right.\\
&\qquad\left.\times\;Z_{k+1}\otimes X_{k+1}|\,\mathcal{F}_{k+1}^{Y}\right]. \qquad (62)
\end{aligned}$$

From the definition of C and $\alpha_{i,j}$ we have

$$\begin{aligned}
&\left\langle C(Z_{k+1}\otimes X_{k+1}),f_j\right\rangle Z_{k+1}\otimes X_{k+1}\\
&\quad=\left\langle C\sum_{i_0=1}^{L}\sum_{j_0=1}^{N}\langle Z_{k+1}\otimes X_{k+1},\alpha_{i_0,j_0}\rangle\alpha_{i_0,j_0},f_j\right\rangle\alpha_{i_0,j_0}\\
&\quad=\sum_{i_0=1}^{L}\sum_{j_0=1}^{N}\left\langle C_{j_0}^{i_0}\langle Z_{k+1}\otimes X_{k+1},\alpha_{i_0,j_0}\rangle,f_j\right\rangle\alpha_{i_0,j_0}.
\end{aligned}$$

Based on this observation, (62) further reduces to

$$\begin{aligned}
q_{k+1}&=M\sum_{j=1}^{M}\langle Y_{k+1},f_j\rangle\overline{E}\\
&\qquad\left[\Lambda_k\sum_{i_0=1}^{L}\sum_{j_0=1}^{N}\left\langle C_{j_0}^{i_0}\left\langle Z_{k+1}\otimes X_{k+1},\alpha_{i_0,j_0}\right\rangle,f_j\right\rangle\right.\\
&\qquad\left.\times\;\alpha_{i_0,j_0}|\,\mathcal{F}_{k+1}^{Y}\right]\\
&=M\sum_{j=1}^{M}\langle Y_{k+1},f_j\rangle\sum_{i_0=1}^{L}\sum_{j_0=1}^{N}c_{j,j_0}^{i_0}\alpha_{i_0,j_0}\\
&\qquad\times\left\langle\overline{E}\left[\Lambda_k Z_{k+1}\otimes X_{k+1}|\mathcal{F}_{k+1}^{Y}\right],\alpha_{i_0,j_0}\right\rangle\\
&=M\sum_{j=1}^{M}\langle Y_{k+1},f_j\rangle\sum_{i_0=1}^{L}\sum_{j_0=1}^{N}c_{j,j_0}^{i_0}\alpha_{i_0,j_0}\\
&\qquad\times\sum_{i_1=1}^{L}\sum_{j_1=1}^{N}\left\langle B_{i_1}\otimes A_{j_1}^{i_1},\alpha_{i_0,j_0}\right\rangle\left\langle q_k,\alpha_{i_1,j_1}\right\rangle \qquad (63)
\end{aligned}$$

where (63) follows from Lemma 3.

Now we write (63) in a compact form. Following the basic properties of Kronecker product, we have:

$$\begin{aligned}
\left\langle B_{i_1}\otimes A_{j_1}^{i_1},\alpha_{i_0,j_0}\right\rangle&=\left\langle B_{i_1}\otimes A_{j_1}^{i_1},h_{i_0}\otimes e_{j_0}\right\rangle\\
&=\left[(B_{i_1})^{\top}\otimes\left(A_{j_1}^{i_1}\right)^{\top}\right]\cdot[h_{i_0}\otimes e_{j_0}]\\
&=\langle B_{i_1},h_{i_0}\rangle\otimes\left\langle A_{j_1}^{i_1},e_{j_0}\right\rangle=b_{i_0,i_1}^{k}a_{j_0,j_1}^{i_1}.
\end{aligned}$$

Therefore

$$\begin{aligned}
&\sum_{i_1=1}^{L}\sum_{j_1=1}^{N}\left\langle B_{i_1}\otimes A_{j_1}^{i_1},\alpha_{i_0,j_0}\right\rangle\left\langle\alpha_{i_1,j_1},q_k\right\rangle\\
&\quad=\sum_{i_1=1}^{L}\sum_{j_1=1}^{N}a_{j_0,j_1}^{i_1}b_{i_0,i_1}^{k}\left\langle\alpha_{i_1,j_1},q_k\right\rangle\\
&\quad=\sum_{i_1=1}^{L}b_{i_0,i_1}^{k}\left\langle h_{i_1}\otimes\left({}_{j_0}A^{i_1}\right)^{\top},q_k\right\rangle\\
&\quad=\sum_{i_1=1}^{L}\left\langle\mathrm{diag}\left\{b_{i_0,i_1}^{k}\otimes\mathbf{1}_N\right\}\left(h_{i_1}\otimes\left({}_{j_0}A^{i_1}\right)^{\top}\right),q_k\right\rangle\\
&\quad=\left\langle\mathrm{diag}\left\{{}_{i_0}B\otimes\mathbf{1}_N\right\}\left({}_{j_0}A\right)^{\top},q_k\right\rangle.
\end{aligned}$$

In this way, we have

$$
\begin{aligned}
q_{k+1} &= M \sum_{j=1}^{M} \langle Y_{k+1}, f_j \rangle \sum_{i_0=1}^{L} \sum_{j_0=1}^{N} c_{j,j_0}^{i_0} \alpha_{i_0,j_0} \\
&\quad \times \sum_{i_1=1}^{L} \sum_{j_1=1}^{N} \left\langle B_{i_1} \otimes A_{j_1}^{i_1}, \alpha_{i_0,j_0} \right\rangle \left\langle \alpha_{i_1,j_1}, q_k \right\rangle \\
&= M \sum_{j=1}^{M} \langle Y_{k+1}, f_j \rangle \sum_{i_0=1}^{L} \sum_{j_0=1}^{N} c_{j,j_0}^{i_0} \alpha_{i_0,j_0} \\
&\quad \times \left\langle \operatorname{diag}\left\{ {}_{i_0}B \otimes \mathbf{1}_N \right\} \left({}_{j_0}A \right)^{\top}, q_k \right\rangle \\
&= \sum_{i_0=1}^{L} \sum_{j_0=1}^{N} \operatorname{diag}\left(M Y_{k+1}^{\top} C \right) \alpha_{i_0,j_0} \cdot \left({}_{j_0}A \right) \\
&\quad \cdot \operatorname{diag}\left\{ {}_{i_0}B \otimes \mathbf{1}_N \right\} q_k \\
&= \operatorname{diag}\left(M Y_{k+1}^{\top} C \right) \sum_{i_0=1}^{L} \sum_{j_0=1}^{N} \alpha_{i_0,j_0} \cdot \left({}_{j_0}A \right) \\
&\quad \cdot \operatorname{diag}\left\{ {}_{i_0}B \otimes \mathbf{1}_N \right\} q_k \\
&= \operatorname{diag}\left(M Y_{k+1}^{\top} C \right) \sum_{i_0=1}^{L} (h_i \otimes A) \cdot \operatorname{diag}\left\{ {}_{i_0}B \otimes \mathbf{1}_N \right\} q_k \\
&= \operatorname{diag}\left(M Y_{k+1}^{\top} C \right) \cdot (I_L \otimes A) \cdot \left[\operatorname{diag}\left\{ {}_{i_0}B \otimes \mathbf{1}_N \right\} \right]_{i_0 \in \mathbb{N}_{1:L}} \cdot q_k \\
&= \operatorname{diag}\left(M Y_{k+1}^{\top} C \right) \cdot (I_L \otimes A) \\
&\quad \cdot \left\{ \left[\operatorname{diag}\left(\operatorname{vec}(B^{\top}) \right) \cdot (\mathbf{1}_L \otimes I_L) \right] \otimes I_N \right\} q_k.
\end{aligned}
$$

This completes the proof of the theorem.

ACKNOWLEDGMENT

The authors would like to thank the Associate Editor and the anonymous reviewers for their suggestions which have improved the quality of the work.

REFERENCES

[1] A. Cardenas, S. Amin, and S. Sastry, "Research challenges for the security of control systems," in *Proc. 3rd Conf. Hot Topics Security*, 2008.

[2] J. Slay and M. Miller, "Lessons learned from the maroochy water breach," in *Critical Infrastructure Protection*, ser. IFIP International Federation for Information Processing. New York, NY, USA: Springer US, 2007, vol. 253, pp. 73–82.

[3] J. Farwell and R.-F. Rohozinski, "StuxNet and the future of cyber war," *Survival*, vol. 53, no. 1, pp. 23–40, 2011.

[4] A.-H. Mohsenian-Rad and A. Leon-Garcia, "Distributed internet-based load altering attacks against smart power grids," *IEEE Trans. Smart Grid*, vol. 2, no. 4, pp. 667–674, Dec. 2011.

[5] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," in *Proc. 16th ACM Conf. Computer Commun. Security*, 2009, pp. 21–32.

[6] Y. Mo and B. Sinopoli, "Secure control against replay attacks," in *Proc. 47th Annu. Allerton Conf. Commun., Control, Computing (Allerton 2009)*, Sep. 2009, pp. 911–918.

[7] Y. Mo, R. Chabukswar, and B. Sinopoli, "Detecting integrity attacks on SCADA systems," *IEEE Trans. Control Syst. Technol.*, vol. 22, no. 4, pp. 1396–1407, Jul. 2014.

[8] F. Miao, M. Pajic, and G. Pappas, "Stochastic game approach for replay attack detection," in *Proc. IEEE 52nd Annu. Conf. Decision Control*, Dec. 2013, pp. 1854–1859.

[9] S. Amin, A. Cardenas, and S. Sastry, "Safe and secure networked control systems under denial-of-service attacks," in *Hybrid Systems: Computation and Control*, ser. Lecture Notes in Computer Science R. Majumdar and P. Tabuada, Eds. Heidelberg, Germany: Springer Berlin Heidelberg, 2009, vol. 5469, pp. 31–45.

[10] H. Zhang, P. Cheng, L. Shi, and J. Chen, "Optimal DoS attack policy against remote state estimation," in *Proc. IEEE 52nd Annu. Conf. Decision Control (CDC)*, Dec. 2013, pp. 5444–5449.

[11] H. Zhang, P. Cheng, L. Shi, and J. Chen, "Optimal denial-of-service attack scheduling with energy constraint," *IEEE Trans. Autom. Control*, vol. 60, no. 11, pp. 3023–3028, Nov. 2015, doi: 10.1109/TAC.2015.2409905.

[12] F. Pasqualetti, F. Dorfler, and F. Bullo, "Attack detection and identification in cyber-physical systems," *IEEE Trans. Autom. Control*, vol. 58, no. 11, pp. 2715–2729, Nov. 2013.

[13] H. Fawzi, P. Tabuada, and S. Diggavi, "Secure estimation and control for cyber-physical systems under adversarial attacks," *IEEE Trans. Autom. Control*, vol. 59, no. 6, pp. 1454–1467, Jun. 2014.

[14] M. Pajic, J. Weimer, N. Bezzo, P. Tabuada, O. Sokolsky, I. Lee, and G. Pappas, "Robustness of attack-resilient state estimators," in *Proc. Int. Conf. Cyber-Physical Syst. (ICCPS)*, Apr. 2014.

[15] Y. Mo, J. Hespanha, and B. Sinopoli, "Resilient detection in the presence of integrity attacks," *IEEE Trans. Signal Process.*, vol. 62, no. 1, pp. 31–43, Jan. 2014.

[16] S. Sundaram, M. Pajic, C. Hadjicostis, R. Mangharam, and G. Pappas, "The wireless control network: Monitoring for malicious behavior," in *Proc. 49th IEEE Conf. Decision Control (CDC)*, Dec. 2010, pp. 5979–5984.

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

[18] F. Pasqualetti, F. Dorfler, and F. Bullo, "Cyber-physical attacks in power networks: Models, fundamental limitations and monitor design," in *Proc. 50th IEEE Conf. Decision Control and Eur. Control Conf. (CDC-ECC)*, Dec. 2011, pp. 2195–2201.

[19] H. Fawzi, P. Tabuada, and S. Diggavi, "Secure state-estimation for dynamical systems under active adversaries," in *Proc. 49th Annu. Allerton Conf. Commun., Control, Computing (Allerton)*, Sep. 2011, pp. 337–344.

[20] H. Zhang, P. Cheng, L. Shi, and J. Chen, "Optimal denial-of-service attack scheduling against linear quadratic Gaussian control," in *Proc. Amer. Control Conf. (ACC)*, Jun. 2014, pp. 3996–4001.

[21] Y. Mo and B. Sinopoli, "Secure estimation in the presence of integrity attacks," *IEEE Trans. Autom. Control*, vol. 60, no. 4, pp. 1145–1151, Apr. 2015.

[22] A. Teixeira, I. Shames, H. Sandberg, and K. H. Johansson, "A secure control framework for resource-limited adversaries," *Automatica*, vol. 51, no. 0, pp. 135–148, 2015.

[23] K. M. Thayer and D. L. Beveridge, "Hidden Markov models from molecular dynamics simulations on DNA," *Proc. Nat. Acad. Sci.*, vol. 99, no. 13, pp. 8642–8647, 2002.

[24] W. Huang, J. Zhang, and Z. Liu, "Activity recognition based on hidden Markov models," in *Knowledge Science, Engineering and Management*, ser. Lecture Notes in Computer Science Z. Zhang and J. Siekmann, Eds. Heidelberg, Germany: Springer Berlin Heidelberg, 2007, vol. 4798, pp. 532–537.

[25] A. Mihoub, G. Bailly, and C. Wolf, "Social behavior modeling based on incremental discrete hidden Markov models," in *Human Behavior Understanding*, ser. Lecture Notes in Computer Science A. Salah, H. Hung, O. Aran, and H. Gunes, Eds. Berlin, Germany: Springer Int., 2013, vol. 8212, pp. 172–183.

[26] J. Yang, Y. Xu, and C. S. Chen, "Human action learning via hidden Markov model," *IEEE Trans. Syst., Man Cybern., A: Syst. Humans*, vol. 27, no. 1, pp. 34–44, Jan. 1997.

[27] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. R. Stat. Soc., Series B*, vol. 39, no. 1, pp. 1–38, 1977.

[28] R. Neal and G. E. Hinton, "A view of the EM algorithm that justifies incremental, sparse, other variants," in *Learning in Graphical Models*. Amsterdam, The Netherlands: Kluwer Academic, 1998, pp. 355–368.

[29] T. Al-ani, "Hidden Markov models in dynamic system modelling and diagnosis," in *Hidden Markov Models, Theory and Applications*, P. Dymarski, Ed. Rijeka, Croatia: InTech, 2011, pp. 27–50.

[30] L. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proc. IEEE*, vol. 77, no. 2, pp. 257–286, Feb. 1989.
[31] M. Gales and S. Young, "The application of hidden Markov models in speech recognition," *Found. Trends Signal Process.*, vol. 1, no. 3, pp. 195–304, 2008.
[32] B. Hannaford and P. Lee, "Hidden Markov model analysis of force/torque information in telemanipulation," *Int. J. Robotics Res.*, vol. 10, no. 5, pp. 528–539, 1991.
[33] S. Zonouz, K. Rogers, R. Berthier, R. Bobba, W. Sanders, and T. Overbye, "SCPSE: Security-oriented cyber-physical state estimation for power grid critical infrastructures," *IEEE Trans. Smart Grid*, vol. 3, no. 4, pp. 1790–1799, Dec. 2012.
[34] O. Geramifard, J. Xu, J. Zhou, and X. Li, "Multimodal hidden Markov model-based approach for tool wear monitoring," *IEEE Trans. Industr. Electron.*, vol. 61, no. 6, pp. 2900–2911, Jun. 2014.
[35] C. McGhan, A. Nasir, and E. Atkins, "Human intent prediction using Markov decision processes," presented at the Infotech@Aerospace 2012 Conf., 2012.
[36] A. H. Tai, W. K. Ching, and L. Y. Chan, "Detection of machine failure: Hidden markov model approach," *Comput. Industr. Eng.*, vol. 57, no. 2, pp. 608–619, 2009.
[37] R. Elliott and W. Malcolm, "Discrete-time expectation maximization algorithms for Markov-modulated poisson processes," *IEEE Trans. Autom. Control*, vol. 53, no. 1, pp. 247–256, Feb. 2008.
[38] R. J. Elliott, L. Aggoun, and J. B. Moore, *Hidden Markov Models: Estimation and Control*. Berlin, Germany: Springer, 1995.
[39] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1979.
[40] B. C. Levy, *Principles of Signal Detection and Parameter Estimation*. Berlin, Germany: Springer, 2008.
[41] Analog Devices, AD7988 Datasheet, 2012. [Online]. Available: http://www.analog.com/static/imported-files/data_sheets/AD7988-1_7988-5.pdf
[42] S. Gillijns and B. D. Moor, "Unbiased minimum-variance input and state estimation for linear discrete-time systems with direct feedthrough," *Automatica*, vol. 43, no. 5, pp. 934–937, 2007.
[43] J. Keller, K. Chabir, and D. Sauter, "Input reconstruction for networked control systems subject to deception attacks and data losses on control signals," *Int. J. Sys. Sci.*, vol. 47, no. 4, pp. 814–820, 2016, doi: 10.1080/00207721.2014.906683.
[44] J. Morris, "The Kalman filter: A robust estimator for some classes of linear quadratic problems," *IEEE Trans. Inform. Theory*, vol. IT-22, no. 5, pp. 526–534, 1976.
[45] Y. Li, L. Shi, P. Cheng, J. Chen, and D. Quevedo, "Jamming attacks on remote state estimation in cyber-physical systems: A game-theoretic approach," *IEEE Trans. Autom. Control*, vol. 60, no. 10, pp. 2831–2836, Oct. 2015, doi: 10.1109/TAC.2015.2461851.

Dawei Shi received the B.Eng. degree in electrical engineering and its automation from the Beijing Institute of Technology, Beijing, China, in 2008, and the Ph.D. degree in control systems from the University of Alberta, Calgary, AB, Canada, in 2014.

Since December 2014, he has been appointed as an Associate Professor at the School of Automation, Beijing Institute of Technology. His research interests include event-based control and estimation, robust model predictive control and tuning, and wireless sensor networks. He is a reviewer for a number of international journals, including the IEEE TRANSACTIONS ON AUTOMATIC CONTROL, *Automatica,* and *Systems and Control Letters*. In 2009, he received the Best Student Paper Award in the IEEE International Conference on Automation and Logistics.

Robert J. Elliott received the bachelors and masters degrees from Oxford University, Oxford, U.K., and the Ph.D. and D.Sc. degrees from Cambridge University, Cambridge, U.K.

He has held positions at Newcastle, Yale, Oxford, Warwick, Hull, Alberta, and visiting positions in Toronto, Northwestern, Kentucky, Brown, Paris, Denmark, Hong Kong, and Australia. From 2001 to 2009 he was the RBC Financial Group Professor of Finance at the University of Calgary, Canada, where he was also an Adjunct Professor in both the Department of Mathematics and the Department of Electrical Engineering. From 2009 to 2014 he was an Australian Professorial Fellow at the University of Adelaide. He has authored nine books and over 470 papers. His book with PE Kopp *Mathematics of Financial Markets* was published by Springer in 1999 and has been reprinted three times. The Hungarian edition was published in 2000 and the second edition was published in September 2004. An edition in China was published in 2010. Springer Verlag published his book *Binomial Methods in Finance*, written with John van der Hoek, in the summer of 2005. He has also worked in signal processing, and his book with L. Aggoun and J. Moore on *Hidden Markov Models: Estimation and Control* was published in 1995 by Springer Verlag and reprinted in 1997. A revised and expanded edition was printed in 2008. His book with L. Aggoun *Measure Theory and Filtering* was published by Cambridge University Press in June 2004. His earlier book *Stochastic Calculus and Applications* was published by Springer in 1982, and a Russian translation appeared in 1986. A greatly enlarged second edition written with S Cohen was published in December 2015.

Tongwen Chen (F'06) received the B.Eng. degree in automation and instrumentation from Tsinghua University, Beijing, China, in 1984, and the M.A.Sc. and Ph.D. degrees in electrical engineering from the University of Toronto, Toronto, ON, Canada, in 1988 and 1991, respectively.

He is presently a Professor of electrical and computer engineering at the University of Alberta, Edmonton, AB, Canada. His research interests include computer and network-based control systems, process safety, and alarm systems, and their applications to the process and power industries. He has served as an Associate Editor for several international journals, including the IEEE TRANSACTIONS ON AUTOMATIC CONTROL, *Automatica*, and *Systems and Control Letters*. He is a Fellow of IFAC, as well as the Canadian Academy of Engineering.