2018 Annual American Control Conference (ACC)
June 27–29, 2018. Wisconsin Center, Milwaukee, USA

State-Secrecy Codes for Stable Systems

Anastasios Tsiamis, Konstantinos Gatsis, George J. Pappas

Abstract— **In this paper, we study the problem of remote state estimation, in the presence of a passive eavesdropper. An authorized user estimates the state of a stable linear plant, based on the packets received from a sensor, while the packets may also be intercepted by the eavesdropper. Our goal is to design a coding scheme at the sensor, which encodes state information, in order to impair the eavesdropper's estimation performance, while enabling the user to successfully decode the sent messages. We introduce a novel class of codes, termed State-Secrecy Codes. By using acknowledgment signals from the user, they apply linear time-varying transformations to the current and previously received states, imposing artificial unstable dynamics to the eavesdropper's Riccati recursive estimation scheme. By exploiting the process noise, the channel randomness and the artificial dynamics, these codes manage to be fast and efficient for real-time dynamical systems. Under minimal conditions they achieve perfect secrecy, namely the eavesdropper's estimation error covariance matrix converges to the open-loop prediction one, which is the same as not intercepting any messages at all; meanwhile, the user's estimation performance is optimal. Those conditions only require that at least once, the user receives the corresponding packet while the eavesdropper fails to intercept it. The theoretical results are illustrated in simulations.**

I. Introduction

The recent emergence of the Internet of Things (IoT) as a collection of interconnected sensors and actuators has created a new attack surface for adversarial attacks [1], [2]. Research efforts in the context of control systems have targeted denial-of-service attacks [3] and data integrity of compromised sensors [4], [5], [6], [7]. Another fundamental vulnerability of such interconnected systems is eavesdropping attacks, especially when the underlying communication medium is of a broadcast nature, i.e. as in wireless systems [8]. This data leakage, not only compromises confidentiality, but could be also used to perform other more complex attacks [9].

In this paper, we study eavesdropping attacks in dynamical systems. In many IoT applications, sensors collect the state information of a dynamical system and send it to an authorized user, i.e. a controller, a cloud server, etc. through a (wireless) channel. Our goal is to design codes such that the user receives the confidential state information, while any eavesdroppers are confused about the true state. Since we are dealing with time-critical systems, it is desirable to avoid elaborate codes which might introduce severe delays to the user's data processing. Thus, we might not be able to employ cryptography-based tools [10], since they might introduce computation and communication overheads [11].

Another approach includes developing codes in the physical layer of wireless communications [12], [13], by exploiting the characteristics of the underlying channel. Information-theoretic approaches, both for static sources [14]–[16] and dynamical systems [17]–[19], give conditions about the existence of codes such that an eavesdropper receives no information. However, finding such codes is challenging in practice and may require knowledge of the eavesdropper's channel, which may not be available. Nonetheless, in the case of packet erasure channels, more practical codes can be designed [20].

A control-theoretic approach was employed in [21]–[23], where the performance metric of the user and the eavesdropper is the minimum mean square error (mmse). These works employ a secrecy mechanism which withholds measurements either randomly [21] of deterministically [22], [23]. Their goal is to achieve large expected mmse error for the eavesdropper while the user's expected mmse error is small. The guarantees about the eavesdropper's error, however, are only in expectation (her error will be small with high probability), while the user's performance is degraded as a side-effect.

In this paper, we develop a novel class of codes, suitable for stable dynamical systems, which we call State-Secrecy Codes. The system's state is encoded by subtracting a weighted version of the user's most recently received state from the current state. This operation has low complexity and only requires acknowledgment signals from the user back to the sensor. By designing the code's weighting factor, we can impose artificial dynamics to the eavesdropper's Riccati recursion, which impair its estimation performance. Then, confidentiality is protected by also exploiting the process noise of the dynamical system and the channel's randomness.

In previous work [24], we developed the first state-secrecy codes for unstable systems, which are usually open-loop control systems. However, the codes from [24] do not work in stable or closed-loop systems, which are much harder to protect; the stable dynamics make any noise to naturally contract over time. In this work, we introduce a new code construction, which imposes artificial unstable dynamics to the eavesdropper's Riccati recursion, counterbalancing the natural stable dynamics.

In Section II we model the dynamical system as linear and the channel as a packet dropping one. We also introduce a novel control-theoretic notion of perfect secrecy, requiring that the eavesdropper's mmse covariance matrix converges to the open-loop prediction one almost surely, while the user's performance is optimal. In Section III, we show that

This work was supported in part by ONR N00014-17-1-2012, and by NSF CNS-1505799 grant and the Intel-NSF Partnership for Cyber-Physical Systems Security and Privacy.

The authors are with the Department of Electrical and Systems Engineering, University of Pennsylvania, Philadelphia, PA 19104. Emails: {atsiamis,kgatsis,pappasg}@seas.upenn.edu

with State-Secrecy Codes, perfect secrecy can be guaranteed under remarkably mild conditions. These conditions require that at some time the user receives the corresponding packet, while the eavesdropper misses it. Due to the artificial dynamics imposed to the eavesdropper's estimation scheme, a single occurrence of this event, which we call critical event, makes the eavesdropper lose track of the state.

In summary our main contributions, are the following:

- We introduce State-Secrecy Codes, which are suitable for stable real-time dynamical systems. They impose artificial dynamics to the eavesdropper's estimation scheme, overcoming the limitations of the codes in [24], which do not work for stable systems.
- Our coding scheme is asymptotically optimal, since the eavesdropper's mmse covariance matrix converges almost surely to the open-loop prediction one, while the user's estimation performance remains optimal. This supersedes the results in [21], [22].
- The condition for perfect secrecy is minimal and channel free, requiring one occurrence of the critical event.

We illustrate the code performance in simulations in Section IV, and conclude with remarks in Section V.

II. Problem formulation

A. Dynamical system model

The considered remote estimation architecture is shown in Figure 1 and consists of a sensor observing a dynamical system, a legitimate user, and an eavesdropper. The dynamical system is linear and has the following form:

$$x_{k+1} = Ax_k + w_{k+1} \tag{1}$$

where $x_k \in \mathbb{R}^n$ is the state, $A \in \mathbb{R}^{n\times n}$ is the system matrix, and $k \in \mathbb{N}$ is the (discrete) time. Signal $w_k \in \mathbb{R}^n$ is the process noise, modeled as i.i.d. Gaussian process with zero mean and covariance matrix Q. The initial state x_0 is also Gaussian with zero mean, covariance Σ_0 and is independent of the process noise. All system and noise parameters A, Q, Σ_0 are assumed to be public knowledge, available to all involved entities, i.e., the sensor, the user, and the eavesdropper. We assume a common probability space Ω for all random quantities (noises, initial condition and channel outcomes). The following assumption holds throughout this paper.

Assumption 1: System (1) is asymptotically stable and matrix A is invertible. Matrices Q, Σ_0 are positive definite. In more compact notation Q, $\Sigma_0 \succ 0$, where $\succ$ ($\succeq$) denotes comparison in the positive definite (semidefinite) cone. $\diamond$

The invertibility assumption is necessary for our coding scheme to work and covers many systems of practical interest. The positive definiteness of Q implies that the process noise can affect all states and create uncertainty about them.

B. Channel model

The sensor communicates over a channel with two outputs/receivers as shown in Figure 1. The input to the channel is denoted by $z_k \in \mathbb{R}^n$. The first output, denoted by $h_{u,k}$, is the authorized one to the user, while the second, denoted by h_k, is the unauthorized one to the eavesdropper. Communication follows the packet-based paradigm commonly used in networked control systems [25]–[27].

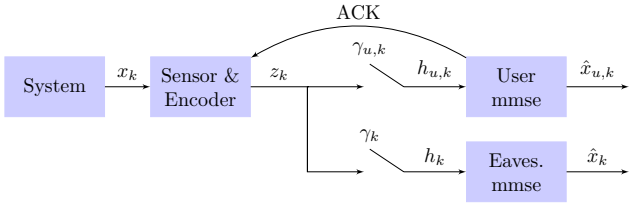

Fig. 1. A sensor collects the state x_k of the dynamical system (1). Then it transmits an encoded version z_k of the state to the channel, which is neither reliably nor securely received by the user. The packets might be dropped, as captured by $\gamma_{u,k}$, and might be intercepted by the eavesdropper, as captured by γ_k. To decode the messages, the user and the eavesdropper use the minimum mean square error (mmse) estimates $\hat{x}_{u,k}$ and $\hat{x}_k$ respectively.

Communication with the user is unreliable, i.e., may undergo packet drops. Additionally, communication is not secure against the eavesdropper, i.e., the latter may intercept transmitted packets. We denote by $\gamma_{u,k} \in \{0, 1\}$ the outcome of the user packet reception at time k, and by $\gamma_k \in \{0, 1\}$ the outcome of the eavesdropper's packet interception. If $\gamma_{u,k} = 1$ (or $\gamma_k = 1$), then the reception (interception) is successful. Otherwise, the respective packet is dropped. The outputs of the channel are modeled as:

$$h_{u,k} = \begin{cases} z_k, & \text{if } \gamma_{u,k} = 1 \\ \varepsilon, & \text{if } \gamma_{u,k} = 0 \end{cases}, h_k = \begin{cases} z_k, & \text{if } \gamma_k = 1 \\ \varepsilon, & \text{if } \gamma_k = 0 \end{cases} \tag{2}$$

where symbol ε, is used to represent the "no information" outcome. The channel outcomes $\{\gamma_{u,k}, \gamma_k, k = 0, 1, \dots\}$ are modeled as random and assumed to be independent of the initial state x_0, and the process noise w_k, for $k = 0, 1, \dots$. We do not assume any specific joint distribution of the channel outcomes; as explained in Section III the result of this paper (Theorem 1) is channel-free.

In addition to the main channel, the user can reliably send acknowledgment signals back to the sensor via the reverse channel. Thus, at any time step the sensor knows what is the latest received message z_k at the user. Meanwhile, we assume that the eavesdropper is able to intercept all acknowledgment signals, and thus, knows the history of user's packet successes. In that respect, we model a powerful eavesdropper. Neither the sensor nor the user have any knowledge about the eavesdropper's intercept successes γ_k.

C. Encoder definition

The sensor collects state measurements x_k and encodes them by sending $z_k \in \mathbb{R}^p$ over the channel at each time step k, where p is an integer to be designed. In general, the encoder may produce z_k, given all the information at the sensor at time k, i.e. current and past states x_t for $t \le k$, past sent messages z_t for $t < k$, as well as past user's channel outcomes $\gamma_{u,t}$ for $t < k$:

$$z_k = f_k\left(x_k, x_t, z_t, \gamma_{u,t}, t < k\right), \tag{3}$$

where f_k is a function from $\mathbb{R}^{m(k+1)+pk} \times \{0,1\}^k$ to $\mathbb{R}^p$. Although this allows encoders with infinite memory, our proposed one (see Section III) does not need the whole history and, thus, only uses finite memory.

D. MMSE Estimation

Both the user and the eavesdropper know the encoding scheme and use a minimum mean square error (mmse) estimate to decode the received/intercepted messages. This estimate depends on their information up to time k. We define the user's batch vector of channel outputs by $\boldsymbol{h}_{u,0:k} = (h_{u,0}, \dots, h_{u,k})$ and the batch vector of channel outcomes by $\boldsymbol{\gamma}_{u,0:k} = (\gamma_{u,0}, \dots, \gamma_{u,k})$. The eavesdropper's batch vectors $\boldsymbol{h}_{0:k}$, $\boldsymbol{\gamma}_{0:k}$ are defined similarly. Then, the user's information at time k is denoted by $\mathcal{I}_k^u = \{\boldsymbol{h}_{u,0:k}\}$, with $\mathcal{I}_{-1}^u = \emptyset$. Respectively, we denote the eavesdropper's information by

$$\mathcal{I}_k = \left\{\boldsymbol{h}_{0:k}, \boldsymbol{\gamma}_{u,0:k}\right\}, \mathcal{I}_{-1} = \emptyset \tag{4}$$

Notice that the eavesdropper has the additional information of the user's reception success history.

The user's mmse estimate, $\hat{x}_{u,k}$, and the respective mmse covariance matrix $P_{u,k}$ are given by:

$$\hat{x}_{u,k} = \mathbb{E}\left\{x_k | \mathcal{I}_k^u\right\}, \quad P_{u,k} = \mathrm{Cov}\left\{x_k | \mathcal{I}_k^u\right\} \tag{5}$$

where the conditional covariance of any random vector Z with respect to some other random vector $\mathcal{I}$ is defined as

$$\mathrm{Cov}\left\{Z|\mathcal{I}\right\} = \mathbb{E}\left\{\left(Z - \mathbb{E}\left\{Z|\mathcal{I}\right\}\right)\left(Z - \mathbb{E}\left\{Z|\mathcal{I}\right\}\right)' |\mathcal{I}\right\}.$$

Similarly, the eavesdropper's mmse estimate, $\hat{x}_k$ and the respective mmse covariance matrix P_k are:

$$\hat{x}_k = \mathbb{E}\left\{x_k | \mathcal{I}_k\right\}, \quad P_k = \mathrm{Cov}\left\{x_k | \mathcal{I}_k\right\}. \tag{6}$$

E. Problem

The goal of this work is to design a coding scheme at the sensor, so that we achieve *perfect secrecy* (introduced in the following definition). We require the user's estimation scheme to be optimal, while the confidentiality guarantees against the eavesdropper are asymptotically optimal. The user's estimation scheme is optimal when at the successful reception times the estimation error is zero. Respectively, the confidentiality guarantees are asymptotically optimal when the eavesdropper's mmse covariance matrix P_k approaches the open-loop prediction one, i.e. the eavesdropper's mmse covariance when all signals z_k are lost. The open-loop prediction estimate and error covariance matrix are given by:

$$x_k^{op} = \mathbb{E}\left\{x_k\right\} = 0, \quad P_k^{op} = \mathrm{Cov}\left\{x_k\right\}, \tag{7}$$

with the prediction error obeying the Lyapunov recursion:

$$P_k^{op} = A P_{k-1}^{op} A' + Q. \tag{8}$$

Definition 1 (Perfect Secrecy): Given the stable system (1) and channel model (2), a coding scheme (3) achieves perfect secrecy if and only if both of the following hold:

(i) the user's performance is optimal:

$$\hat{x}_{u,k} = x_k, \text{ when } \gamma_{u,k} = 1. \tag{9}$$

(ii) the eavesdropper's mmse covariance matrix converges to the open-loop prediction error covariance matrix with probability one:

$$P_k - P_k^{op} \overset{a.s.}{\to} 0, \tag{10}$$

where $\overset{a.s.}{\to}$ denotes almost sure convergence with respect to the probability space Ω, as $k \to \infty$ and the convergence is with respect to any matrix norm. $\diamond$

Remark 1: Contrary to the unstable case in [24], [28] where we required the trace $\mathbf{tr}\, P_k$ to converge to infinity, here we require the whole covariance matrix P_k to converge to the open-loop prediction covariance matrix. Here the objective is different since we require that the whole state is protected.

III. State-Secrecy Codes for Stable Systems

In this section, we introduce State-Secrecy Codes for stable systems, which, under minimal conditions, lead to perfect secrecy. The sensor encodes and transmits the current state x_k as a weighted state difference of the form $x_k - L^{k-t_k} x_{t_k}$, where L is a carefully designed matrix and x_{t_k} is a previous state called the *reference state* of the encoded message, for some $t_k < k$ depending on k. The sensor and the user can agree on this reference state via the acknowledgment signals, e.g., it can be the most recent state received at the user's end. At the user's side no information is lost with this encoding; upon receiving a new message $x_k - L^{k-t_k} x_{t_k}$, she can first recover x_k by adding $L^{k-t_k} x_{t_k}$ and then notify the sensor to use x_k as the reference state for the next transmission.

On the other hand, when the eavesdropper fails to intercept that reference packet at time t_k, we will show that the eavesdropper's mmse covariance matrix starts converging to the open-loop prediction one. That is because the eavesdropper misses important information about the reference state x_{t_k}. Then, by carefully designing L, the eavesdropper's error is amplified when she tries to decode a following packet of the form $x_k - L^{k-t_k} x_{t_k}$ to obtain x_k. This, in turn, also obstructs the eavesdropper from decoding future packets, as any following reference state x_k for some $k > t_k$, depends on the current reference state x_{t_k} and so on. This triggers an irreversible chain reaction effect; the uncertainty about x_{t_k} is amplified through time, driving the eavesdropper's mmse covariance matrix to the open-loop prediction one. For this reason, we call the event, where the user receives a packet at time t_k while the eavesdropper misses it, *critical event*.

Let us now formally present the coding scheme. We define the *reference time* t_k to be the time of the most recent successful reception at the user before k:

$$t_k = \max\left\{t : \ 0 \le t < k, \ \gamma_{u,t} = 1\right\}. \tag{11}$$

When the set $\left\{t : \ 0 \le t < k, \ \gamma_{u,t} = 1\right\}$ is empty (before the first successful transmission), we use $t_k = -1$, $x_{-1} = 0$.

Since by Assumption 1 A is stable and $Q \succ 0$, the open-loop prediction covariance matrix P_k^{op} converges to a steady-state matrix P_L, the unique positive definite solution of:

$$P_L = A P_L A' + Q. \tag{12}$$

Since the solution is positive definite, P_L^{-1} exists.

Algorithm 1 State-Secrecy Code for stable systems

Input: A, Q, x_k for all $k \geq 0$.

Output: Encoded signals z_k, for all $k \geq 0$.

Let t represent the time of user's most recent message.

Solve $P_L = AP_LA' + Q$, set $L = P_L(P_LA')^{-1}$

Initialize $t = -1$, $x_{-1} = 0$

for $k = 0, 1, \ldots$ **do**

  Transmit $z_k = x_k - L^{k-t}x_t$

  if Acknowledgment received **then** $t = k$

  end if

end for

Definition 2 (State-Secrecy Codes for stable systems): Given system (1) and under Assumption 1, a State-Secrecy Code applies the following time-varying linear operation

$$z_k = x_k - L^{k-t_k}x_{t_k}, \text{ with } L = P_L(P_LA')^{-1}, \qquad (13)$$

where t_k is the reference time as defined in (11) and P_L is the solution of the Lyapunov equation (12). $\diamond$

This choice of L imposes unstable dynamics to the eavesdropper's estimation as explained in Remark 2. Since matrices A, P_L are invertible, matrix L in (13) is well-defined.

The implementation of the scheme is described in Algorithm 1. The sensor always keeps in memory the reference time t_k and state x_{t_k}, with $t_0 = -1$, $x_{-1} = 0$. At each time step k, it transmits z_k as in (13). If the user receives the packet successfully, it sends an acknowledgment signal back to the sensor. In this case, the sensor updates the reference time $t_{k+1} = k$. Otherwise, it keeps $t_{k+1} = t_k$. The memory required for the encoder is minimal ($\mathcal{O}(n)$) and the only computational burden is a matrix-vector multiplication ($\mathcal{O}(n^2)$). The critical event described in the beginning of the section and formally defined below, is crucial for reinforcing secrecy with our coding scheme.

Definition 3 (Critical event): A critical event occurs at time k if the user receives the packet, while the eavesdropper fails to intercept it: $\gamma_{u,k} = 1$, $\gamma_k = 0$ $\diamond$

An example to clarify the coding scheme and the critical event is presented next.

Example 1: Suppose that for $k = 0, 1, 2, 3$ we have the channel outcomes as shown in the first two rows of the following table:

k	0	1	2	3
user $\gamma_{u,k}$	0	1	0	1
eavs. γ_k	1	0	0	1
t_k	-1	-1	1	1
z_k	x_0	x_1	$x_2 - Lx_1$	$x_3 - L^2x_1$
user $h_{u,k}$	ε	x_1	ε	$x_3 - L^2x_1$
eavs. h_k	x_0	ε	ε	$x_3 - L^2x_1$

Then, the last four rows of the table are constructed using the definitions of the reference times (11), of the coding scheme (13), and the channel outcomes (2). Notice that a critical event occurs at time $k = 1$, when the user receives x_1, while the eavesdropper misses it. Then, the user can recover x_3 at time $k = 3$, adding L^2x_1 to $h_{u,3}$. However, since the eavesdropper does not know x_1, she cannot recover x_3. Since $\gamma_{u,3} = 1$, x_3 is going to be the next reference state after $k = 3$. Thus, the eavesdropper will also not be able to successfully recover x_k, for $k > 3$. A single occurrence of the critical event impairs future estimation at the eavesdropper. $\diamond$

The following theorem, formally proves the previous observations. If the critical event $\{\gamma_{u,k_0} = 1, \gamma_{k_0} = 0\}$ occurs at some time k_0, then the eavesdropper's mmse covariance matrix starts to converge to the open-loop prediction one. On the other hand, the user's performance remains optimal.

Theorem 1 (Perfect secrecy): Consider system (1), with channel model (2) and coding scheme (13). If

$$\mathbb{P}(\gamma_{u,k} = 1, \gamma_k = 0, \text{ for some } k \geq 0) = 1, \qquad (14)$$

then perfect secrecy is achieved according to Definition 1. $\diamond$

The condition (14) for perfect secrecy is minimal; it only requires the critical event, where the user receives a message without the eavesdropper intercepting it, to occur at least once. Any joint distribution of packet receptions and interceptions that satisfies this condition is covered. In this sense, the result is channel-free, and holds in most cases of practical interest (see also Remark 1 in [24]).

The proof of Theorem 1 is included in the Appendix and is a consequence of the following lemma, which can be thought as the worst case, in terms of secrecy, of Theorem 1. That is when the critical event $\{\gamma_{u,k_0} = 1, \gamma_{k_0} = 0\}$ occurs at time k_0 and the eavesdropper receives all the packets for $k > k_0$.

Lemma 1 (Worst case analysis): Consider system (1) with channel model (2) and coding scheme (13). Suppose

$$\mathcal{B} = \{\gamma_{u,k_0} = 1, \gamma_{k_0} = 0\} \text{ and} \qquad (15)$$

$$\mathcal{C} = \{\gamma_k = 1, \text{ for all } k \geq k_0 + 1\} \qquad (16)$$

both occur for some $k_0 \geq 0$. Then, conditioned on $\mathcal{B} \cap \mathcal{C}$:

1) For $k \geq k_0$, the eavesdropper's mmse covariance matrix satisfies the Riccati recursion:

$$P_{k+1} = AP_kA' + Q - T_k\left(HP_kH' + Q\right)^{-1}T_k', \qquad (17)$$

where $T_k = AP_kH' + Q$, and $H = A - L$.

2) The mmse covariance matrix P_k converges to P_L:

$$P_k \to P_L, \qquad (18)$$

where P_L is defined in (12). $\diamond$

The above result shows that even in the most pessimistic case of Theorem 1 for confidentiality, the eavesdropper's mmse covariance matrix converges to the open-loop prediction one. In the general case when the eavesdropper does not intercept all packets after k_0, her mmse covariance will be larger (cf. proof of Theorem 1 in Appendix).

The proof of (18) is based on the following lemma. The Riccati recursion (17) is different from the standard one, since we have the additive term of Q in T_k. It still can be transformed to a standard Riccati recursion (see equation (19) below), which, however, violates the stabilization condition. Nonetheless, using the results from [29], we prove that P_k converges to P_L under the assumption that the initial condition is positive definite.

Lemma 2 (Riccati Convergence): Let A, Q satisfy Assumption 1. Suppose P_k satisfies equation (17) for $k \geq k_0$, with L defined in (13), P_L defined in (12) and $H = A - L$. Then, the following properties hold:

1) Recursion (17) is equivalent to:

$$\begin{aligned} &P_{k+1} = g\left(P_k\right), \\ &g(X) = LXL' - LXH'\left(HXH' + Q\right)^{-1} HXL' \end{aligned} \tag{19}$$

2) The pair (L, H) is observable.
3) P_L is the stabilizing solution of $P_L = g(P_L)$ (see [29] for definition).
4) If $P_{k_0} \succ 0$, then $P_k \to P_L$ exponentially fast. ◇

The rate of the exponential convergence in the above lemma is asymptotically equal to

$$P_k - P_L \sim A(P_{k-1} - P_L)A',$$

as follows from the proof of Lemma 2 (part 3) in the Appendix and equation (4.2) in [29]. The following remark reveals the intuition behind the form of matrix L.

Remark 2 (Artificial Dynamics): Comparing (19) with the standard Riccati recursion:

$$P_{k+1} = \bar{A}P_k\bar{A}' + \bar{Q} - \bar{A}P_k\bar{C}'\left(\bar{C}P_k\bar{C}' + \bar{R}\right)^{-1}\bar{C}P_k\bar{A}',$$

we have $\bar{A} = L$, $\bar{C} = A - L$, $\bar{Q} = 0$ and $\bar{R} = Q$. Thus, by selecting an unstable L, we impose unstable artificial dynamics to the eavesdropper's estimation scheme to counterbalance the stable dynamics. Thus, we make the mmse covariance matrix in (19) to converge to a positive definite solution if $P_{k_0} \succ 0$; if L was stable then (19) would converge to zero. The specific selection of L in (13) tunes the steady-state solution of (19) to be equal to P_L. It forces the encoded messages z_k to be less and less correlated with x_k for the eavesdropper, with $\mathrm{Cov}\left\{z_k, x_k | \mathcal{I}_{k-1}\right\} = 0$, when $P_{k-1} = P_L$.

Remark 3: One caveat is that the first time k_0 of the critical event is random and not in our control. A possible remedy is to use a more expensive coding scheme, i.e. encryption, to securely and reliably transmit just the first packet at time $k = 0$. Then, letting our cheap coding scheme take over is sufficient to achieve perfect secrecy. ◇

IV. Simulations

In this section, we illustrate the efficiency of State-Secrecy Codes via numerical simulation. The system under consideration has state matrix $A = \begin{bmatrix} 0.9 & 1 \\ 0 & 0.8 \end{bmatrix}$ and noise covariance matrices $\Sigma_0 = Q = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 2 \end{bmatrix}$. For the channel model, we assume that the channel outcomes are independent across time and stationary with probabilities $P(\gamma_{u,k} = i, \gamma_k = j) = p_{ij}$, for $i, j \in \{0, 1\}$. The assumed values are $p_{11} = 0.7$, $p_{01} = p_{10} = p_{00} = 0.1$. For the estimation scheme of the eavesdropper we used equation (20) in the Appendix (see [28] for more details). Since the user can decode all signals, we used the formula:

$$P_{u,k} = \begin{cases} 0 & \text{if } \gamma_{u,k} = 1 \\ AP_{u,k-1}A' + Q & \text{if } \gamma_{u,k} = 0 \end{cases}$$

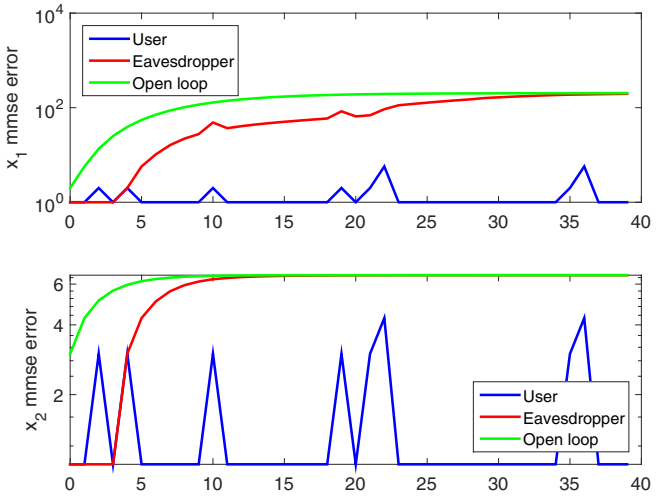

Fig. 2. We compare the eavesdropper's, user's and open-loop diagonal covariance entries for the states x_1 and x_2. For the log-plots we used function $\log(x+1)$ instead of $\log x$. For this random sample of channel outcomes, the critical event occurs at time $k_0 = 5$. After $k = 5$ the eavesdropper's error starts converging to the open-loop prediction one for both sub-states. The user has zero error at the reception times.

In Figure 2, we plot the user's and eavesdropper's mmse errors for the states x_1, x_2, i.e. the diagonal elements of the covariance matrices $P_{u,k}$, P_k. We compare them to the open-loop prediction error P_k^{op} defined in (7). As shown in Figure 2, the eavesdropper's mmse error starts converging to the open-loop prediction one after the first critical event occurs at time $k_0 = 5$. The user can decode all received messages and has zero error at the times of successful reception.

V. Conclusion

The presence of the eavesdropper adds new challenges to the problem of remote estimation. Nonetheless, by using a simple State-Secrecy Code, based on acknowledgment signals from the user back to the sensor, we can achieve powerful confidentiality guarantees with minimal computational cost. At the same time the user's estimation performance is optimal. By exploiting the imposed artificial dynamics, the random packet erasures and the process noise, perfect secrecy is achieved with just a single occurrence of the critical event, when the user receives more information than the eavesdropper. Future work includes implementation and experimental evaluation of the proposed scheme. We also seek to adapt our codes to the case of output measurements. Finally, the analysis should be repeated when the eavesdropper uses a different estimator than the mmse one.

Appendix

A. Estimation formula

In the following proposition, we present the estimation formula for the eavesdropper's mmse covariance. The proof is similar to [28] and hence, omitted.

Proposition 1 ([28]): Consider system (1), with channel (2) and coding scheme (13). Fix any $k \geq 0$. Let the

covariance matrix of x_k and z_k given $\mathcal{I}_{k-1}$ be:

$$\mathrm{Cov}\left\{ \begin{bmatrix} x_k \\ z_k \end{bmatrix} |\mathcal{I}_{k-1} \right\} = \begin{bmatrix} \Sigma_{k,xx} & \Sigma_{k,xz} \\ \Sigma_{k,zx} & \Sigma_{k,zz} \end{bmatrix}$$

Then, the eavesdropper's mmse covariance P_k at time k defined in (6) is given by:

$$P_k = \Sigma_{k,xx} - \gamma_k \Sigma_{k,xz} \left(\Sigma_{k,zz}\right)^{-1} \Sigma_{k,zx}, \tag{20}$$

Moreover, $\Sigma_{k,xx} = AP_{k-1}A' + Q$ if $k > 0$ and $\Sigma_{0,xx} = \Sigma_0$.

B. Proof of Lemma 1

First, we prove (17). By independence of w_{k+1} from $\mathcal{I}_k$, it follows that $x_{k+1} - \mathbb{E}\{x_{k+1}|\mathcal{I}_k\} = A(x_k - \hat{x}_k) + w_{k+1}$. Next, we claim that in $\mathcal{B} \cap \mathcal{C}$, for $k \geq k_0$:

$$z_{k+1} - \mathbb{E}\{z_{k+1}|\mathcal{I}_k\} = (A - L)(x_k - \hat{x}_k) + w_{k+1}, \tag{21}$$

which we prove in the next paragraph. Thus, denoting $H = A - L$, in $\mathcal{B} \cap \mathcal{C}$ we have:

$$\begin{aligned} &\Sigma_{k+1,xx} = AP_kA' + Q,\ \Sigma_{k+1,xz} = AP_kH' + Q, \\ &\Sigma_{k+1,zz} = HP_kH' + Q. \end{aligned} \tag{22}$$

Thus, formula (17) follows from (20), (22).

Next, we show (21). Since the critical event happened at time k_0, the reference time at k_0+1 is $t_{k_0+1} = k_0$ (equation (11)). Hence, all reference times t_{k+1} , for $k \geq k_0$ satisfy $t_{k+1} \geq k_0$ and there are two possible cases:

1) $t_{k+1} = k$, when $\gamma_{u,k} = 1$
2) $t_{k+1} = t_k$, when $\gamma_{u,k} = 0$.

In the former one, the intercepted signal by (13) is $z_{k+1} = x_{k+1} - Lx_k = (A-L)x_k + w_{k+1}$. But the process noise w_{k+1} is independent of $\mathcal{I}_k$, thus, $\mathbb{E}\{w_{k+1}|\mathcal{I}_k\} = 0$ and equation (21) holds. In the latter one, we have

$$z_{k+1} = x_{k+1} - L^{k+1-t_{k+1}} x_{t_{k+1}} = x_{k+1} - L^{k+1-t_k} x_{t_k}.$$

since $t_{k+1} = t_k$. Then, if we add and subtract Lx_k, we get:

$$\begin{aligned} z_{k+1} &= x_{k+1} - Lx_k + L(x_k - L^{k-t_k} x_{t_k}). \\ &= (A - L)x_k + w_{k+1} + Lz_k \end{aligned} \tag{23}$$

But since $\gamma_{u,k} = 0$ and $t_k \geq k_0$, the only possibility is $k > k_0$. Thus, the eavesdropper has intercepted z_k, which implies $Lz_k = \mathbb{E}\{Lz_k|\mathcal{I}_k\}$ in $\mathcal{B} \cap \mathcal{C}$ or

$$\mathbb{E}\{z_{k+1}|\mathcal{I}_k\} = (A - L)\hat{x}_k + Lz_k.$$

This, along with (23) prove equation (21).

Convergence (18) follows from Lemma 2, if we show that $P_{k_0} \succ 0$. But since $\gamma_{k_0} = 0$, from (20), we either have $P_{k_0} = \Sigma_0 \succ 0$ or $P_{k_0} = AP_{k_0-1}A' + Q \succeq Q \succ 0$ ■

C. Technical lemmas

Lemma 3 (Covariance upper bound): Consider system (1), with channel model (2) and coding scheme (13). Then, with probability one:

$$P_k \preceq P_k^{op}, \text{ for all } k \geq 0, \tag{24}$$

where P_k is the eavesdropper's mmse covariance matrix (6) and P_k^{op} is the open-loop prediction covariance matrix (7).

Proof: We use induction. For $k = 0$, from estimation formula (20):

$$P_0 \preceq \Sigma_0 = P_0^{op}.$$

Next, assume that the induction hypothesis $P_k \preceq P_k^{op}$ is true. Then, for $k+1$, we have:

$$P_{k+1} \preceq AP_kA' + Q \preceq AP_k^{op}A' + Q = P_{k+1}^{op},$$

where the first inequality follows from (20) and the second from the induction hypothesis and monotonicity of operator $AXA' + Q$ with respect to X. ■

Lemma 4 (Kalman filter with correlated noise. [30]): Let $\bar{A}, \bar{Q} \in \mathbb{R}^{n\times n}$, $\bar{S}', \bar{C} \in \mathbb{R}^{m\times n}$ and $\bar{R} \in \mathbb{R}^{m\times m}$, for some $n, m \in \mathbb{N}$, $m \leq n$. If

$$\begin{bmatrix} \bar{Q} & \bar{S} \\ \bar{S}' & \bar{R} \end{bmatrix} \succeq 0,\ \bar{R} \succ 0,$$

then, the Kalman recursion

$$\bar{P}_{k+1} = \bar{A}\bar{P}_k\bar{A}' + \bar{Q} - \bar{T}_k \left(\bar{C}\bar{P}_k\bar{C}' + \bar{R}\right)^{-1} \bar{T}_k', \tag{25}$$

with $\bar{T}_k = \bar{A}\bar{P}_k\bar{C}' + \bar{S}$ is equivalent to:

$$\bar{P}_{k+1} = A_s\bar{P}_kA_s' + Q_s - A_s\bar{P}_k\bar{C}' \left(\bar{C}\bar{P}_k\bar{C}' + \bar{R}\right)^{-1} \bar{C}\bar{P}_kA_s',$$

with $A_s = \bar{A} - \bar{S}\bar{R}^{-1}\bar{C}$, $Q_s = \bar{Q} - \bar{S}\bar{R}^{-1}\bar{S}'$. ◇

D. Proof of Lemma 2

To prove 1), we apply Lemma 4. Notice that equation (17) has the form of (25) with $\bar{A} = A$, $\bar{C} = H$, $\bar{Q} = \bar{S} = \bar{R} = Q$ and covariance matrix:

$$\begin{bmatrix} Q & Q \\ Q & Q \end{bmatrix} = \begin{bmatrix} Q^{1/2} \\ Q^{1/2} \end{bmatrix} \begin{bmatrix} Q^{1/2} & Q^{1/2} \end{bmatrix} \succeq 0$$

Since $Q \succ 0$, equation (17) is equivalent to (19) as follows from the transformations of Lemma 4 with $A_s = L$, $Q_s = 0$.

To prove 2), assume that (L, H) is not observable. By the PHB observability conditions, there exists an eigenvector v of L such that $Lv = \lambda v$ and $Hv = (A - L)v = 0$. But this implies that also $Av = \lambda v$. This is a contradiction since A is a stable matrix, while $L = P_L(A')^{-1}(P_L)^{-1}$ has only unstable eigenvalues. Thus, (L, H) is observable.

To prove 3) first notice that if $P_k = P_L$, then in (17):

$$T_k = AP_L(A' - L') + Q = P_L - AP_LL' = 0.$$

Thus, $P_{k+1} = AP_LA' + Q = P_L = P_k$ and P_L is a fixed point of (17). From 1), it is also a fixed point of (19). It remains to show that P_L is a stabilizing solution (see [29]). This requires the matrix $L - K_\infty H$, to have all eigenvalues inside the unit circle, where

$$K_\infty = LP_LH' \left(HP_LH' + Q\right)^{-1}.$$

Applying the matrix inversion lemma, after some lengthy algebra we find that $K_\infty = -I$. Thus, $L - K_\infty H = A$. Since A is asymptotically stable, P_L is stabilizing.

Part 4) follows directly from Theorem 4.2 in [29]. Since L does not have any eigenvalue on the unit circle, (L, H) is observable and $P_{k_0} \succ 0$, P_k converges to the unique stabilizing solution P_L exponentially fast. ■

E. Proof of Theorem 1

First, we show that if the critical event $\mathcal{B} = \{\gamma_{u,k_0} = 1, \gamma_{k_0} = 0\}$ occurs at some time k_0, then the eavesdropper's error covariance matrix converges to P_L. Define a new channel outcome sequence $\tilde{\gamma}_{u,k}$, $\tilde{\gamma}_k$ that is coupled with the original outcome sequence $\gamma_{u,k}$, γ_k as follows:

$$(\tilde{\gamma}_{u,k}, \tilde{\gamma}_k) = (\gamma_{u,k}, \gamma_k), \text{ for all } 0 \leq k \leq k_0$$
$$(\tilde{\gamma}_{u,k}, \tilde{\gamma}_k) = (\gamma_{u,k}, 1), \text{ for all } k > k_0.$$

In this new outcome sequence, the eavesdropper receives everything after time k_0. Now, define again the channel outputs (2), the eavesdropper's information (4) and the covariance matrix (6), but with the original channel outcomes $\gamma_{u,k}$, γ_k replaced by $\tilde{\gamma}_{u,k}$, $\tilde{\gamma}_k$, for all k. Denote the eavesdropper's covariance matrix under $\tilde{\gamma}_{u,k}$, $\tilde{\gamma}_k$ by $\tilde{P}_k$. Following the proof of Lemma 5 in [28] we obtain the inequality $P_k \succeq \tilde{P}_k$ with probability one. But from the above inequality and Lemma 3:

$$\tilde{P}_k \preceq P_k \preceq P_k^{op}, \quad \text{with prob. 1}, \tag{26}$$

where P_k^{op} is the open-loop prediction covariance matrix defined in (8), which converges to P_L. Thus, it is sufficient to show that also $\tilde{P}_k$ converges in $\mathcal{B}$. Notice that since $\tilde{\gamma}_{u,k_0} = \gamma_{u,k_0}$ and $\tilde{\gamma}_{k_0} = \gamma_{k_0}$, the following events are equal:

$$\tilde{\mathcal{B}} = \{\tilde{\gamma}_{u,k_0} = 1, \tilde{\gamma}_{k_0} = 0\} = \mathcal{B}.$$

Also observe that since $\tilde{\gamma}_{2,k} = 1$, $k > k_0$, the event

$$\tilde{\mathcal{C}} = \{\tilde{\gamma}_{u,k} = 1, \text{ for all } k \geq k_0 + 1\}$$

is the whole probability space. Thus, applying Lemma 1 for the coupled channel sequence $\tilde{\gamma}_{u,k}$, $\tilde{\gamma}_k$, and the events $\tilde{\mathcal{B}}$ and $\tilde{\mathcal{C}}$, we obtain $\tilde{P}_k \to P_L$ in $\mathcal{B}$. Hence, from (26) we get:

$$P_k \to P_L, \text{ in } \mathcal{B}.$$

Finally, we prove that coding scheme (13) achieves perfect secrecy. The above convergence result along with the theorem hypothesis (14), prove that:

$$P_k \to P_L, \text{ with prob. 1}.$$

But since also $P_k^{op} \to P_L$, we prove condition (10) of perfect secrecy. The user always knows x_{t_k} and, thus, she can completely reconstruct the states x_k, when $\gamma_{u,k} = 1$. This exactly implies that condition (9) is satisfied. ■

REFERENCES

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

[2] H. Sandberg, S. Amin, and Johansson, K.H. (Organizers), "Cyberphysical Security in Networked Control Systems [Special Issue]," *IEEE Control Systems*, vol. 35, no. 1, 2015.

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

[4] M. Pajic, J. Weimer, N. Bezzo, P. Tabuada, O. Sokolsky, I. Lee, and G. J. Pappas, "Robustness of attack-resilient state estimators," in *ICCPS'14: ACM/IEEE 5th International Conference on Cyber-Physical Systems (with CPS Week 2014)*. IEEE Computer Society, 2014, pp. 163–174.

[5] Y. Mo, J. P. Hespanha, and B. Sinopoli, "Resilient detection in the presence of integrity attacks," *IEEE Transactions on Signal Processing*, vol. 62, no. 1, pp. 31–43, 2014.

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

[7] F. Pasqualetti, F. Dorfler, and F. Bullo, "Control-theoretic methods for cyberphysical security: Geometric principles for optimal cross-layer resilient control systems," *IEEE Control Systems*, vol. 35, no. 1, pp. 110–127, 2015.

[8] Y. Zou, J. Zhu, X. Wang, and L. Hanzo, "A survey on wireless security: Technical challenges, recent advances, and future trends," *Proceedings of the IEEE*, vol. 104, no. 9, pp. 1727–1765, Sept 2016.

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

[10] J. Katz and Y. Lindell, *Introduction to modern cryptography*. CRC press, 2014.

[11] J. Lee, K. Kapitanova, and S. H. Son, "The price of security in wireless sensor networks," *Computer Networks*, vol. 54, no. 17, pp. 2967–2978, 2010.

[12] P. A. Regalia, A. Khisti, Y. Liang, and Tomasin, S. (Eds.), "Secure Communications via Physical-Layer and Information-Theoretic Techniques [Special Issue]," *Proceedings of the IEEE*, vol. 103, no. 10, 2015.

[13] M. Bloch, J. Barros, M. R. Rodrigues, and S. W. McLaughlin, "Wireless information-theoretic security," *IEEE Transactions on Information Theory*, vol. 54, no. 6, pp. 2515–2534, 2008.

[14] A. D. Wyner, "The wire-tap channel," *The Bell System Technical Journal*, vol. 54, no. 8, pp. 1355–1387, Oct 1975.

[15] Y. Liang, H. V. Poor, and S. Shamai, "Secure communication over fading channels," *IEEE Transactions on Information Theory*, vol. 54, no. 6, pp. 2470–2492, 2008.

[16] F. Oggier and B. Hassibi, "The secrecy capacity of the mimo wiretap channel," *IEEE Transactions on Information Theory*, vol. 57, no. 8, pp. 4961–4972, Aug 2011.

[17] H. Li, L. Lai, and W. Zhang, "Communication requirement for reliable and secure state estimation and control in smart grid," *IEEE Transactions on Smart Grid*, vol. 2, no. 3, pp. 476–486, 2011.

[18] M. Wiese, K. H. Johansson, T. J. Oechtering, P. Papadimitratos, H. Sandberg, and M. Skoglund, "Uncertain wiretap channels and secure estimation," in *2016 IEEE International Symposium on Information Theory (ISIT)*, July 2016, pp. 2004–2008.

[19] ——, "Secure estimation for unstable systems," in *IEEE 55th Conference on Decision and Control (CDC)*. IEEE, 2016, pp. 5059–5064.

[20] I. Safaka, L. Czap, K. Argyraki, and C. Fragouli, "Creating secrets out of packet erasures," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 6, pp. 1177–1191, 2016.

[21] A. Tsiamis, K. Gatsis, and G. J. Pappas, "State estimation with secrecy against eavesdroppers," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 8385–8392, 2017.

[22] A. S. Leong, D. E. Quevedo, D. Dolz, and S. Dey, "On remote state estimation in the presence of an eavesdropper," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 7339–7344, 2017.

[23] A. S. Leong, D. E. Quevedo, and S. Dey, "State estimation over markovian packet dropping links in the presence of an eavesdropper," in *IEEE 56th Conference on Decision and Control (CDC)*, 2017.

[24] A. Tsiamis, K. Gatsis, and G. J. Pappas, "State estimation codes for perfect secrecy," in *IEEE 56th Conference on Decision and Control (CDC)*, 2017.

[25] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. I. Jordan, and S. S. Sastry, "Kalman filtering with intermittent observations," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1453–1464, 2004.

[26] J. P. Hespanha, P. Naghshtabrizi, and Y. Xu, "A survey of recent results in networked control systems," *Proceedings of the IEEE*, vol. 95, no. 1, p. 138, 2007.

[27] K. Gatsis, A. Ribeiro, and G. J. Pappas, "Optimal power management in wireless control systems," *IEEE Transactions on Automatic Control*, vol. 59, no. 6, pp. 1495–1510, 2014.

[28] A. Tsiamis, K. Gatsis, and G. J. Pappas, "State-secrecy codes for networked linear systems," 2017, arXiv preprint arXiv:1709.04530.

[29] S. Chan, G. Goodwin, and K. Sin, "Convergence properties of the Riccati difference equation in optimal filtering of nonstabilizable systems," *IEEE Transactions on Automatic Control*, vol. 29, no. 2, pp. 110–118, 1984.

[30] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear estimation*. Prentice Hall, 2000.