2017 IEEE 56th Annual Conference on Decision and Control (CDC)
December 12-15, 2017, Melbourne, Australia

# State Estimation Codes for Perfect Secrecy

Anastasios Tsiamis, Konstantinos Gatsis, George J. Pappas

*Abstract*— **We study the problem of remote state estimation, in the presence of a passive eavesdropper. An authorized user estimates the state of an unstable 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 the 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, suitable for real-time dynamical systems, termed State-Secrecy Codes. By using acknowledgment signals from the user, they apply linear time-varying transformations to the current and previously received states. We prove that under minimal conditions, State-Secrecy Codes achieve perfect secrecy, namely the eavesdropper's minimum mean square error grows unbounded almost surely, while the user's estimation performance is optimal. These conditions 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 tempting front for cyber-attacks [1], [2]. Eavesdropping attacks, which compromise confidentiality of information, are a fundamental vulnerability of such interconnected systems, especially when the underlying medium of communication is of a broadcast nature as in wireless systems [3]. In this paper, we study eavesdropping attacks in the context of real-time dynamical systems. In many IoT applications, sensors collect state information about 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 data processing of the user. Thus, we might not be able to employ cryptography-based tools [4], which are mostly used in practice, since they might introduce computation and communication overheads [5].

Another approach, includes developing codes in the physical layer of wireless communications [6]. This approach exploits the characteristics of the underlying communication channel, e.g., the wireless medium. Information-theoretic tools are used [7]–[9] to 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 [10]. Although the aforementioned approaches typically involve static sources, recently, in [11]–[13], information-theoretic tools were used in the case of dynamical systems, in remote estimation scenarios. Still those approaches face the same challenges with the static ones.

A control-theoretic approach was employed in [14], [15], where the performance metric of the user and the eavesdropper is the minimum mean square error (mmse). Instead of encoding, these works employ a secrecy mechanism which withholds measurements either randomly [14] or deterministically [15]. In the case of unstable systems, under certain conditions, the eavesdropper's expected mmse error can grow to infinity, while the user's expected mmse error remains bounded. However, the main disadvantage is that the guarantees about the eavesdropper's error are only in expectation, not almost surely, while the user's performance is degraded as a side-effect.

In this paper, we develop a novel class of codes, suitable for unstable real-time 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. To protect confidentiality, State-Secrecy codes exploit the inherent process noise of the dynamical system, the channel's randomness, and the system's dynamics.

In Section II we model the dynamical system as linear and the channel as a packet dropping one. Similar to [12]–[14] we assume that the system is unstable. We also introduce a novel control-theoretic notion of secrecy, requiring that the eavesdropper's mmse error grows unbounded almost surely, while the user's performance is optimal. In Section III, we show that 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. Just a single occurrence of this event, which we call critical event, makes the eavesdropper lose track of the system state. In summary, our main contributions are the following:

- We introduce State-Secrecy Codes, which are suitable for real-time dynamical systems. Their efficiency does not depend on the eavesdropper's computational capabilities.

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

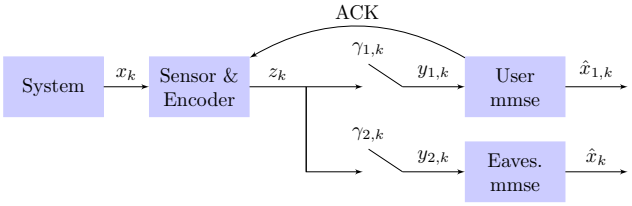


Fig. 1. A sensor collects the state $x_k$ of the dynamical system (1). Then it transmits an encoded version $z_k$ of the state to the channel, which is neither reliably nor securely received by the user. The packets might be dropped, as captured by $\gamma_{1,k}$, and might be intercepted by the eavesdropper, as captured by $\gamma_{2,k}$. To decode the messages, the user and the eavesdropper use the minimum mean square error (mmse) estimates $\hat{x}_{1,k}$ and $\hat{x}_k$ respectively.

- The codes achieve perfect secrecy, i.e. unbounded eavesdropper's mmse error almost surely, and optimal user's estimation performance.
- The condition for perfect secrecy is remarkably minimal, requiring just a single occurrence of the critical event. This condition is also channel free.

We conclude this paper by validating the performance of the secrecy mechanism in simulations in Section IV, and with remarks in Section V.

Due to space limitations some of the proofs are omitted. Full proofs can be found in the extended version of this paper [16], which also contains new results.

## II. Problem formulation

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 and is modeled as an independent Gaussian random variable with zero mean and covariance matrix $Q$. The initial state $x_0$ is also a Gaussian random variable with zero mean and covariance $\Sigma_0$. Matrices $Q$, $\Sigma_0$ are assumed to be positive definite. In more compact notation, $Q$, $\Sigma_0 \succ 0$, where $\succ$ ($\succeq$) denotes comparison in the positive definite (semidefinite) cone. We assume that all system and noise parameters $A, Q, \Sigma_0$ are public knowledge, available to all involved entities, i.e., the sensor, the user, and the eavesdropper. A more general formulation where the sensor observes a system output is also possible (see [16]).

In this paper, we consider an unstable system, i.e., its spectral radius is $\rho(A) = \max_i|\lambda_i(A)| > 1$. From a security point of view, we can achieve much better confidentiality when the system is unstable; without any measurements, the unstable dynamics amplify the uncertainty caused by the process noise. On the other hand, when the system is stable, the problem is more challenging. The eavesdropper can always predict that a stable system is close to equilibrium without even eavesdropping. We do not deal with stable systems in this paper, but it is subject of ongoing work.

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 $y_{1,k}$, is the authorized one to the user, while the second, denoted by $y_{2,k}$, is the unauthorized one to the eavesdropper. Communication follows the packet-based paradigm commonly used in networked control systems [17], [18]. We assume that the packets consist of sufficiently large number of bits to neglect quantization effects [17]–[19].

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. In particular, we denote by $\gamma_{1,k} \in \{0, 1\}$ the outcome of the user packet reception at time $k$, and by $\gamma_{2,k} \in \{0, 1\}$ the outcome of the eavesdropper's packet interception. When $\gamma_{1,k} = 1$ (or $\gamma_{2,k} = 1$), then the reception (interception) is successful. Otherwise the reception (interception) is not successful and the respective packet is dropped. Thus, the outputs of the channel are modeled as:

$$y_{i,k} = \begin{cases} z_k, & \text{if } \gamma_{i,k} = 1 \\ \varepsilon, & \text{if } \gamma_{i,k} = 0 \end{cases} \tag{2}$$

for $i = 1, 2$, where symbol $\varepsilon$, is used to represent the "no information" outcome. The random variables $\{\gamma_{1,k}, \gamma_{2,k}, k = 0, 1, \dots\}$ are assumed independent of the system state $x_k$, $k = 0, 1, \dots$. We do not assume any specific joint distribution of the channel outcomes.

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 $\gamma_{2,k}$.

The sensor collects the state measurements $x_k$ and encodes them by sending $z_k \in \mathbb{R}^p$ over the channel, where $p$ is an integer to be designed. The encoder may produce $z_k$, based on 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 $\gamma_{1,t}$ for $t < k$.

Both the user and the eavesdropper know the coding 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 batch vector of received channel outputs $\boldsymbol{y}_{i,0:k} = (y_{i,0}, \dots, y_{i,k})$ and channel outcomes $\boldsymbol{\gamma}_{i,0:k} = (\gamma_{i,0}, \dots, \gamma_{i,k})$, for $i = 1, 2$. Then, the eavesdropper's information at time $k$ is denoted by

$$\mathcal{I}_k = \left\{\boldsymbol{y}_{2,0:k}, \boldsymbol{\gamma}_{1,0:k}\right\}, \mathcal{I}_{-1} = \emptyset \tag{3}$$

Respectively, we denote the user's information at time $k$ by $\mathcal{I}^1_k = \{\boldsymbol{y}_{1,0:k}\}$, with $\mathcal{I}^1_{-1} = \emptyset$. Notice that the eavesdropper has the additional information of the user's reception success history. With those definitions, the user's mmse estimate,

$\hat{x}_{1,k}$ and the respective mmse covariance matrix $P_{1,k}$ are:

$$\hat{x}_{1,k} = \mathbb{E}\left\{x_k|\mathcal{I}_k^1\right\}, \quad P_{1,k} = \mathrm{Cov}\left\{x_k|\mathcal{I}_k^1\right\} \tag{4}$$

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 covariance matrix $P_k$ are:

$$\hat{x}_k = \mathbb{E}\left\{x_k|\mathcal{I}_k\right\}, \quad P_k = \mathrm{Cov}\left\{x_k|\mathcal{I}_k\right\}. \tag{5}$$

Now, we can formally state the goal of the paper, which is to design a coding scheme that achieves *perfect secrecy*, introduced in the following definition. We require the eavesdropper's mmse error to grow unbounded, while the user successfully decodes the information and has optimal estimation performance. The estimation performance is optimal if at the successful reception times, the mmse error is the same as if no packet had been dropped (see also [16]).

*Definition 1 (Perfect Secrecy):* Given system (1) and channel model (2), we say that a coding scheme achieves perfect secrecy if and only if the following conditions hold:

(i) the user's performance is optimal:

$$\hat{x}_{1,k} = x_k, \text{ when } \gamma_{1,k} = 1. \tag{6}$$

(ii) the eavesdropper's mmse error diverges to infinity with probability one:

$$\mathbf{tr}\, P_k \overset{a.s.}{\to} \infty, \tag{7}$$

where $\mathbf{tr}$ is the trace operator.

This notion of secrecy is asymptotic; the eavesdropper can maintain a trivial open-loop prediction estimate, i.e. $\hat{x}_k = 0$, that has unbounded but finite estimation error at any time $k$. Moreover, we remark that (7) guarantees aggregate state secrecy in that, at least one but not necessarily all eigenvalues of the eavesdropper's error covariance grow unbounded, e.g., she might still be able to estimate a stable part of the state with bounded error (see also Section IV).

In the following section, we introduce a coding scheme that achieves perfect secrecy by exploiting the unstable system dynamics, the process noise, the acknowledgment signals, as well as a minimal assumption on the channel.

## III. State-Secrecy Codes for Unstable Systems

In this section, we present State-Secrecy Codes, which, along with a very mild sufficient condition on the channel outcomes, lead to perfect secrecy. The sensor encodes and transmits the current state measurement $x_k$ as a weighted state difference of the form $x_k - A^{k-t_k}x_{t_k}$, where $x_{t_k}$ 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 $x_k - A^{k-t_k}x_{t_k}$, she can first recover $x_k$ by adding $A^{k-t_k}x_{t_k}$ and then notify the sensor to use $x_k$ as the reference state for the next transmission.

On the other hand, on the event that the eavesdropper fails to intercept that reference packet at time $t_k$, her error starts increasing. That is because the eavesdropper misses the reference state $x_{t_k}$ and, thus, cannot decode a following packet of the form $x_k - A^{k-t_k}x_{t_k}$ to obtain $x_k$. But this 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, which combined with the instability of the system, leads to an exponentially growing eavesdropper's estimation error. For this reason, we call this event, where the user receives a packet at time $t_k$ while the eavesdropper misses it, the *critical event*.

The following definitions formally describe our coding scheme. We define the *reference time* $t_k$ to be the time of the most recent successful reception at the user before $k$:

$$t_k = \max\left\{0 \le t < k : \gamma_{1,k} = 1\right\}. \tag{8}$$

Until the first successful transmission, i.e. when the set $\max\left\{0 \le t < k : \gamma_{1,t} = 1\right\}$ is empty, we use the convention $t_k = -1$, with $x_{-1} = 0$.

*Definition 2 (State-Secrecy Codes for unstable systems):* Given the unstable system matrix $A$ in (1), a State-Secrecy Code applies the following time-varying linear operation

$$z_k = x_k - A^{k-t_k}x_{t_k}, \tag{9}$$

where $t_k$ is the reference time defined in (8). ◇

The coding scheme is described in Algorithm 1. The sensor only needs finite memory, it only stores the reference time $t_k$ and state $x_{t_k}$, with $x_{-1} = 0$. After each successful reception, the user sends an acknowledgment signal back to the sensor. In this case, the sensor updates the reference time $t_{k+1} = k$. Otherwise, it keeps $t_{k+1} = t_k$. 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 three rows of the following table:

| $k$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| user $\gamma_{1,k}$ | 0 | 1 | 1 | 1 |
| eavs. $\gamma_{2,k}$ | 1 | 0 | 1 | 1 |
| $t_k$ | $-1$ | $-1$ | 1 | 2 |
| $z_k$ | $x_0$ | $x_1$ | $x_2 - Ax_1$ | $x_3 - Ax_2$ |
| user $y_{1,k}$ | $\varepsilon$ | $x_1$ | $x_2 - Ax_1$ | $x_3 - Ax_2$ |
| eavs. $y_{2,k}$ | $x_0$ | $\varepsilon$ | $x_2 - Ax_1$ | $x_3 - Ax_2$ |

Then, the last four rows of the table are constructed using the definitions of the reference times (8), of the coding scheme (9), and the channel outcomes (2). Notice that the critical event occurs at time $k = 1$, when the user receives $x_1$, while the eavesdropper misses it. Then, the user can recover $x_2$ at time $k = 2$, adding $Ax_1$ to $y_{1,2}$. However, since the eavesdropper does not know $x_1$, she cannot recover $x_2$. Since $\gamma_{1,2} = 1$, $x_2$ is the next reference state at time $k = 3$. Thus, the eavesdropper will also not be able to recover $x_3$, from

**Algorithm 1** State-Secrecy Code

**Input:** $A$, $x_k$ at each $k \geq 0$

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

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

Initialize $t = -1$, $x_{-1} = 0$
**for** $k = 0, 1, \ldots$ **do**
  Transmit $z_k = x_k - A^{k-t} x_t$
  **if** Acknowledgment received **then** $t = k$
  **end if**
**end for**

$y_{2,3} = x_3 - Ax_2$. Hence, a single occurrence of the critical event impairs future estimation at the eavesdropper. ◇

Our main result, presented in following theorem, formally proves the previous observations. If the critical event $\{\gamma_{1,k_0} = 1, \gamma_{2,k_0} = 0\}$ occurs at some time $k_0$, then the eavesdropper's error starts to grow unbounded exponentially fast. On the other hand, the user's performance is optimal.

*Theorem 1 (Perfect secrecy):* Consider system (1), with channel model (2) and coding scheme (9). If

$$\mathbb{P}(\gamma_{1,k} = 1, \gamma_{2,k} = 0, \text{ for some } k \geq 0) = 1, \tag{10}$$

then:

(i) perfect secrecy is achieved according to Definition 1.
(ii) conditioned on the event $\{\gamma_{1,k_0} = 1, \gamma_{2,k_0} = 0\}$ for some $k_0 \geq 0$, the eavesdropper's mmse error grows unbounded for $k \geq k_0$ as

$$\mathbf{tr}\, P_k = c\rho\,(A)^{2(k-k_0)} \tag{11}$$

where $P_k$ is the mmse covariance defined in (5) and $c > 0$ is a constant independent of $k_0$. ◇

The above theorem is remarkable as the condition (10) for perfect secrecy is completely 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, and in this sense the result is channel-free, and holds in cases of practical interest – see Remark 1.

The intuition behind the weighting factor $A^{k-t_k}$ in the code is as follows. The difference $x_k - A^{k-t_k} x_{t_k}$, is actually a linear combination of the process noise from time $t_k + 1$ up to $k$:

$$x_k - A^{k-t_k} x_{t_k} = \sum_{j=t_k+1}^{k} A^{k-j} w_j$$

as follows from the system dynamics (1). If the critical event occurs at time $k_0$, then the eavesdropper permanently loses all information about the process noise $w_{k_0}$ at time $k_0$. This loss of information is amplified by the unstable system dynamics over time leading to unbounded error.

The detailed proof of Theorem 1 can be found in [16]. It 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 $\{\gamma_{1,k_0} = 1, \gamma_{2,k_0} = 0\}$ occurs at time $k_0$ and the eavesdropper receives all the following packets for $k > k_0$.

*Lemma 1 (Worst case analysis):* Consider system (1), with channel model (2) and coding scheme (9). If both events:

$$\mathcal{B} = \{\gamma_{1,k_0} = 1, \gamma_{2,k_0} = 0\}$$
$$\mathcal{C} = \{\gamma_{2,k} = 1, \text{ for all } k \geq k_0 + 1\}$$

occur for some $k_0 \geq 0$, then

$$P_k = A^{k-k_0} P_{k_0} \left(A'\right)^{k-k_0}, \tag{12}$$

for $k \geq k_0$ in $\mathcal{B} \cap \mathcal{C}$. ◇

The proof is included in the Appendix. Notice that the equation $P_k = A^{k-k_0} P_{k_0} \left(A'\right)^{k-k_0}$ is unstable with rate $\rho\,(A)^2$. Hence, even in the most pessimistic case of Theorem 1, the eavesdropper still has unbounded error. In the general case when the eavesdropper does not intercept all packets after $k_0$, her mmse error will be even larger (see [16]), which verifies Theorem 1.

*Remark 1:* Suppose that the channel outcomes are independent over time, and suppose that there is a positive probability that the critical event occurs at any time $k$, i.e., $P\,(\gamma_{1,k} = 1, \gamma_{2,k} = 0) > \delta > 0$. For example, in a wireless communication setting this may be due to attenuation of the transmitted signal at the eavesdropper or due to environmental interference. Then, the condition (10) for perfect secrecy of Theorem 1 holds as follows from the Borel-Cantelli lemma [20]. ◇

*Remark 2:* One caveat of our coding scheme is that the first time $k_0$ the critical event occurs, as in the statement (ii) of Theorem 1, is in general random and not in our control. If the eavesdropper's interception rate is very high, the event may take some time to occur. To overcome this problem, we could use a more expensive defense mechanism, i.e. encryption, to force the critical event. In this case, it is sufficient to securely and reliably transmit just the first packet at time $k = 0$. Then, letting our proposed cheap coding scheme take over is sufficient to achieve perfect secrecy. Hence, Theorem 1 allows us to concentrate the more expensive defense efforts in just a single transmission.

*Remark 3:* In our previous work [14] we considered the architecture without acknowledgments; in order for the user to have smaller mmse error than the eavesdropper in expectation, we required the user's reception rate to be higher than the eavesdropper's ($P\,(\gamma_{1,k} = 1) > P\,(\gamma_{2,k} = 1)$). In contrast, here we do not need such a channel disparity requirement. As a result, acknowledgments are of paramount value for secrecy. We note that acknowledgments were used in [15] to decide whether to transmit or not unencoded state information. Here we exploit the acknowledgments for encoding, achieving this way the almost sure guarantees of Theorem 1, which are stronger than the guarantees in expectation in [14], [15]. This comparison is visualized in Section IV, via simulation. ◇

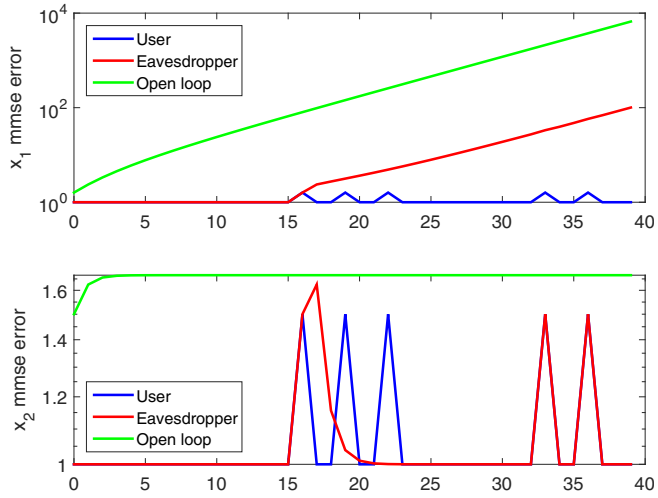


Fig. 2. We compare the eavesdropper's, user's and open-loop mmse errors for the states $x_1$ and $x_2$. For the log-plots, we use function $\log(x+1)$ instead of $\log(x)$. Notice that the critical event occurs at time $k = 17$. The eavesdropper error regarding the unstable part grows unbounded with the same rate as the open loop error. The eavesdropper still has knowledge about the stable stae $x_2$. However, the open-loop prediction error for $x_2$ shows that the eavesdropper has bounded error with respect to the stable dynamics, regardless the defense mechanism.

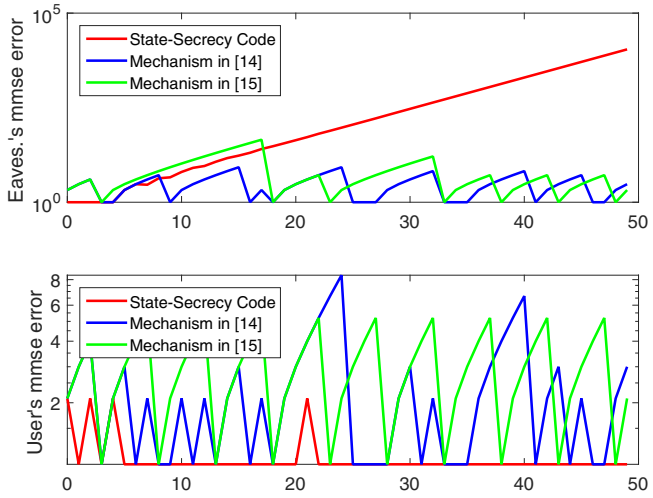


Fig. 3. We compare State-Secrecy Codes with the mechanisms in [14], [15] for a typical channel outcome sequence. For the log-plots, we use function $\log(x+1)$ instead of $\log(x)$. We observe that it significantly outperforms the mechanisms in both confidentiality (eavesdropper's error) and efficiency (user's error).

## IV. Simulations

In this section, we illustrate the efficiency of out coding scheme in numerical simulations. We consider two scenarios. In the first one, we compare the user's and eavesdropper's estimation performance. In the second one, we contrast our results with those achieved by the mechanisms in [14], [15]. The system under consideration has state matrix $A = \begin{bmatrix} 1.2 & 0.1 \\ 0 & 0.5 \end{bmatrix}$ and the noise covariance matrices are $\Sigma_0 = Q = \begin{bmatrix} 0.6 & 0.2 \\ 0.2 & 0.5 \end{bmatrix}$. For the channel model, we assume that the channel outcomes are independent across time and stationary with probabilities $P(\gamma_{1,k} = i, \gamma_{2,k} = j) = p_{ij}$, for $i, j \in \{0, 1\}$. For the estimation scheme of the eavesdropper we used equation (13) in the Appendix (see [16] for details). Since the user can decode all signals, we used the formula:

$$P_{1,k} = (1 - \gamma_{1,k})(AP_{1,k-1}A' + Q)$$

For the first scenario we assume that the channel outcomes have the probabilities $p_{11} = 0.7$, $p_{00} = 0.1$, $p_{01} = 0.1$ and $p_{10} = 0.1$. In Figure 2, we plot the user's and eavesdropper's mmse errors over time for the states $x_1$ and $x_2$ i.e. the two diagonal elements of the covariance matrices respectively. For comparison we also plot the open-loop prediction error, namely the error when all packets are lost. In this case, the open-loop error covariance matrix $P_{o,k}$ is always equal to $P_{o,k} = AP_{o,k-1}A' + Q$. As shown in Figure 2, the eavesdropper's mmse error for the unstable state $x_1$ starts growing unbounded at time $k = 17$, when the critical event occurs. It is worth noting that for state $x_1$, the eavesdropper's rate of increase is the same as in the open loop case. However, the error for state $x_2$ remains bounded, i.e., the system state is only partially protected, since $x_2$ corresponds to a stable system eigenvalue. Nonetheless, it is fundamentally impossible to have unbounded error for the stable part, regardless the defense mechanism, as even the open-loop prediction error is bounded for state $x_2$.

In the second scenario, we assumed $p_{11} = 0.54$, $p_{00} = 0.04$, $p_{01} = 0.06$ and $p_{10} = 0.36$. We compare the performance of our coding scheme with the one of the mechanisms in [14] and [15]–see Figure 3. The comparison is made for the same sequence of channel outcomes, with respect to the user's and eavesdropper's mmse errors $\mathbf{tr}(P_k)$ and $\mathbf{tr}(P_{1,k})$. For the mechanism in [14], which randomly withholds state information with probability $p$, we selected $p = 0.29$. For the infinite horizon mechanism in [15], which transmits state information only if the user loses more than $s$ consecutive packets, we used $s = 5$. Notice that the eavesdropper's mmse error is small very often for the mechanisms in [14] and [15] since unboundedness is guaranteed in expectation, not almost surely. In contrast, our code, achieves unbounded eavesdropper's mmse error almost surely. Also notice that the user's estimation performance is degraded in [14], [15].

## V. Conclusion

The presence of an eavesdropper adds new challenges to the problem of remote estimation. Nonetheless, if the system is unstable, by employing a State-Secrecy Code, based on acknowledgment signals from the user back to the sensor, we can achieve powerful confidentiality guarantees with minimal computational cost. By exploiting the packet erasures, the process noise, and the dynamics, 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 an implementation and experimental evaluation of the proposed scheme. Another open question is

how to adapt our coding scheme to offer more confidentiality guarantees for the stable part of the state.

## APPENDIX

### A. Estimation error covariance formula

In the following proposition, we present the estimation formula for the eavesdropper's mmse covariance. For a detailed proof, one may refer to [16].

*Proposition 1 ( [16]):* Consider system (1), with channel (2) and coding scheme (9). Fix any $k \geq 0$. Let the covariance matrix of $x_k$ and $z_k$ given $\mathcal{I}_{k-1}$ be written in a block form:

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

Then, $\Sigma_{k,xx} = AP_{k-1}A' + Q$ if $k > 0$ and $\Sigma_{0,xx} = \Sigma_0$, where $P_{k-1}$ is the mmse covariance of the eavesdropper at time $k-1$ defined in (5). The mmse covariance at time $k$ is given by:

$$P_k = \Sigma_{k,xx} - \gamma_{2,k}\Sigma_{k,xz}\left(\Sigma_{k,zz}\right)^{\dagger}\Sigma_{k,zx}, \qquad (13)$$

where $(\cdot)^{\dagger}$ denotes the Moore-Penrose pseudoinverse. $\diamond$

### B. Proof of Lemma 1

We will use induction to prove formula (12) of the lemma. For $k = k_0$ it is immediate. Suppose it is true for $k-1 > k_0$. Since in $\mathcal{C}$ the eavesdropper receives all packets for $k > k_0$, we have $\gamma_{2,k} = 1$, for $k > k_0$. According to the recursive formula (13), to find $P_k$, we should compute the covariance of $x_k$ and $z_k$, conditioned on $\mathcal{I}_{k-1}$. By independence of $w_k$ from $\mathcal{I}_{k-1}$, it follows that $x_k - \mathbb{E}\{x_k|\mathcal{I}_{k-1}\} = A\left(x_{k-1} - \hat{x}_{k-1}\right) + w_k$. Now, we claim $z_k - \mathbb{E}\{z_k|\mathcal{I}_{k-1}\} = w_k$, for $k > k_0$, which we prove in the next paragraph. Thus, the covariance matrix of $x_k$ and $z_k$, conditioned on $\mathcal{I}_{k-1}$ is:

$$\mathrm{Cov}\left\{ \left[ \begin{array}{c} x_k \\ z_k \end{array} \right] |\mathcal{I}_{k-1} \right\} = \left[ \begin{array}{cc} AP_{k-1}A' + Q & Q \\ Q & Q \end{array} \right], \text{ in } \mathcal{B}\cap\mathcal{C}$$

Thus, by (13), for $k > k_0$ we have

$$P_k = AP_{k-1}A' + Q - QQ^{-1}Q = AP_{k-1}A' \qquad (14)$$

in $\mathcal{B}\cap\mathcal{C}$. Hence, by the induction hypothesis, we get (12).

Finally, we prove the claim

$$z_k - \mathbb{E}\{z_k|\mathcal{I}_{k-1}\} = w_k, \text{ for } k > k_0, \text{ in } \mathcal{B}\cap\mathcal{C}. \qquad (15)$$

Since the critical event happened at time $k_0$, the reference time at $k_0+1$ is $t_{k_0+1} = k_0$ (equation (8)). Hence, all reference times $t_k$ , for $k > k_0$ satisfy $t_k \geq k_0$ and there are only two possible cases depending on $\gamma_{1,k-1}$:
*Case I:* $t_k = k-1 \geq k_0$, when $\gamma_{1,k-1} = 1$
*Case II:* $t_k = t_{k-1} \geq k_0$ when $\gamma_{1,k-1} = 0$
In the former one, the intercepted signal by (9) is $z_k = x_k - Ax_{k-1} = w_k$. But the process noise $w_k$ is independent of $\mathcal{I}_{k-1}$, thus, $\mathbb{E}\{w_k|\mathcal{I}_{k-1}\} = 0$ and equation (15) holds. In the latter one, we have

$$z_k = x_k - A^{k-t_k}x_{t_k} = x_k - A^{k-t_{k-1}}x_{t_{k-1}}, \qquad (16)$$

since $t_k = t_{k-1}$. Adding and subtracting $Ax_{k-1}$ at the right hand side of the above equation, we obtain

$$\begin{aligned} z_k &= x_k - Ax_{k-1} + A\left(x_{k-1} - A^{k-1-t_{k-1}}x_{t_{k-1}}\right) \\ &= w_k + Az_{k-1}, \end{aligned} \qquad (17)$$

where the second equality follows from the definition of $z_{k-1}$ in (9). But $k-1 > k_0$ (since $\gamma_{1,k-1} = 0$), thus, the eavesdropper has intercepted $z_{k-1}$, which in turn implies $z_{k-1} = \mathbb{E}\{z_{k-1}|\mathcal{I}_{k-1}\}$ in $\mathcal{B}\cap\mathcal{C}$. Hence, from (17), $\mathbb{E}\{z_k|\mathcal{I}_{k-1}\} = Az_{k-1}$ in $\mathcal{B}\cap\mathcal{C}$, which along with (17) prove equation (15). ■

## 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] 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.

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

[5] 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.

[6] 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.

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

[8] 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.

[9] 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.

[10] 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.

[11] 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.

[12] 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.

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

[14] A. Tsiamis, K. Gatsis, and G. J. Pappas, "State estimation with secrecy against eavesdroppers," in *IFAC World Congress*, 2017, to appear. arXiv preprint arXiv:1612.04942.

[15] A. S. Leong, D. E. Quevedo, D. Dolz, and S. Dey, "Remote state estimation over packet dropping links in the presence of an eavesdropper," *arXiv preprint arXiv:1702.02785*, 2017.

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

[17] 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.

[18] 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.

[19] 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.

[20] R. Durrett, *Probability: theory and examples*. Cambridge university press, 2010.