2016 IEEE 55th Conference on Decision and Control (CDC)
ARIA Resort & Casino
December 12-14, 2016, Las Vegas, USA

Secure Estimation for Unstable Systems

Moritz Wiese, Karl Henrik Johansson, Tobias J. Oechtering,
Panos Papadimitratos, Henrik Sandberg, Mikael Skoglund

Abstract— **We assume that the states of an unstable dynamical system are encoded and sent to an estimator through an uncertain channel, which is a channel disturbing its inputs in a nonstochastic manner. The encoder's codeword is also fed into another uncertain channel, at whose output an eavesdropper is listening. The estimator should obtain a uniformly bounded system state estimation error, whereas the eavesdropper's information about the system states should be subject to a security constraint. We find a condition on the relation between the uncertain channel from encoder to estimator and the uncertain channel from encoder to eavesdropper which if satisfied allows for at least one of the two following security criteria to hold: The volume of the set of states possible acccording to the eavesdropper's information tends to infinity at exponential speed—by itself (strong security) or if divided by the volume of the set of states possible according to the estimator's information (weak security).**

I. INTRODUCTION

With the increasing deployment and growing importance of cyber-physical systems, the question of their security has recently become a focus of research activity in control theory [5]. One central vulnerability of networked control or estimation is the communication channel between plant and controller/estimator and possibly the feedback channel from the controller to the plant. One type of attack on the channels is to actively interfere with transmitted information with the goal of degrading the control or estimation performance (e. g. [1], [2], [4], [10], [11]). However, remote estimation of an unstable plant also entails the possibility of eavesdropping. An adversary might have the chance to overhear the transmitted information and to obtain its own system estimate. This paper addresses the question of what can be done to protect the transmitted information from such an attacker.

We consider a scalar, linear, discrete-time, time-invariant linear system. Its initial state is uncertain, i.e., it is an arbitrary element of a bounded interval. The possible initial states are not weighted in any way. Further, the system dynamics is also subject to nonstochastic disturbances which are time-invariant and can assume any value in a fixed, bounded interval. This kind of disturbances is common in robust control.

An estimator has the task of estimating the system state. It produces one estimate for every state produced by the system. Thus for every time step, one can calculate the absolute difference between the state at that time and its estimate. The estimator's goal is that the estimation error, i.e., the supremum over time of all these absolute differences, be bounded uniformly in all possible system state trajectories.

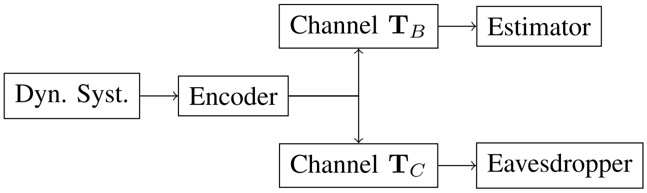

Fig. 1. The problem setup. The pair $(\mathbf{T}_B, \mathbf{T}_C)$ of uncertain channels forms an uncertain wiretap channel.

The estimator does not have direct access to the system state. Instead, an entity called the encoder observes the system state. It is linked to the estimator through an uncertain channel, which has its own finite input and output alphabets. In an uncertain channel, every input generates an output, but these outputs are disturbed in a nonstochastic manner. The encoder transforms every observation into one among a fixed number of messages. This message is mapped to a codeword and fed into the channel. If the encoder uses the right set of codewords, the estimator is able to exactly recover the message transmitted by the encoder, and thus obtains information about the system state.

The codeword sent by the encoder to the estimator is also fed into another, different, uncertain channel, whose output is observed by an adversarial party, the eavesdropper. Since the eavesdropper is assumed to know the encoder's scheme of transforming system states into codewords, in this way it also obtains some information about the system state. The goal is to give the eavesdropper as little information as possible. In mathematical terms, the volume of the set of system states which are possible according to the eavesdropper's information should tend to infinity at exponential speed, at least if divided by the volume of the set of system states which are possible according to the estimator's information. In the former case, we call the security achieved *strong security*, in the latter case *weak security*.

The pair of uncertain channels consisting of the uncertain channel from the encoder to the estimator and the uncertain channel from the encoder to the eavesdropper is called an uncertain wiretap channel. In this paper we identify a class of uncertain wiretap channels which allow for a transmission scheme between the encoder and the estimator such that both of the goals mentioned above are achieved at the same time: the estimator's estimation error is uniformly bounded and the system state is strongly or weakly secure with respect to the eavesdropper. Whether only weak or even strong security

The authors are with the ACCESS Linnaeus Centre, KTH Royal Institute of Technology, SE-10044 Stockholm, Sweden {moritzw, kallej, oech, papadim, hsan, skoglund}@kth.se

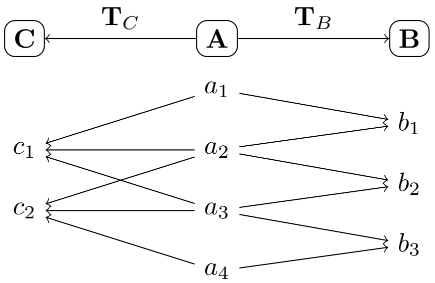

Fig. 2. An uncertain wiretap channel. A line between a_i and b_j indicates that $b_j \in \mathbf{T}_B(a_i)$, similar for a_i and c_j.

is possible depends on how the system parameters compare with the channel parameters. Strong security implies that the eavesdropper's estimation error tends to infinity as time tends to infinity. Central to achieving strong security is the system's instability, which ensures that small uncertainties induced for the eavesdropper by the employed transmission scheme enlarge over time and make the volume of the set of possible states diverge exponentially.

This work is a direct extension of our previous paper [13]. In [13], the setting was exactly the same, only the security criterion differed from the two applied here. In [13], it was required that for every possible sequence of outputs of the uncertain channel from encoder to eavesdropper, there be at least two system state trajectories such that the absolute distance between the two states at a given time grows to infinity as time tends to infinity. This security criterion is weaker than the above strong security criterion, but it is not in general weaker or stronger than the above weak security criterion.

To our knowledge, the only previous paper combining estimation and security for an unstable system is [7], which however considers stochastic disturbances both in the system and the channel and uses a non-operational security criterion based on entropy whose implications are not immediately clear. The motivation for the nonstochastic setup considered in this work comes from Nair [9], who considered estimation and control of an unstable dynamical system with nonstochastic disturbances over an uncertain channel. Nair's restriction of the channel to have nonstochastic disturbances in contrast to channels with stochastic disturbances can be justified with the work [8] by Matveev and Savkin, who proved that if the system and channel disturbances are stochastic and the estimator's goal is to obtain an almost surely bounded estimation error, what matters about the communication channel is only which inputs can generate which outputs, but not with which probability this happens.

The outline of this paper is as follows. In Section II, the problem is defined. Section III gives the main results and proofs of the main theorems, together with an outline of the analysis behind these proofs.

Notation: Logarithms are to the base 2. Given two sets $A, B \subset \mathbb{R}$ and a number λ, we define

$$\lambda A + B := \{c \in \mathbb{R} : c = \lambda a + b \text{ for some } a \in A, b \in B\}.$$

Sequences $(a_t)_{t=0}^T$ of real numbers or integers are denoted by $a(0:T)$, where T may also be infinite.

II. REMOTE ESTIMATION AND UNCERTAIN CHANNELS

Let $\lambda > 0$ and consider the unstable system

$$x(t+1) = \lambda x(t) + w(t). \tag{1}$$

The sequence of disturbances $w(0:\infty)$ is unknown a priori and can be any element in $[-\Omega/2, \Omega/2]^\infty$, where $\Omega \geq 0$ is fixed and known. The initial state $x(0)$ can attain any value in the bounded interval I_0 and is also unknown a priori.

The estimator can only obtain its information from the encoder by transmission through an uncertain channel, and the eavesdropper also listens through an uncertain channel. The problem setup is illustrated in Fig. 1.

Definition 1:
1) An *uncertain channel* is a mapping $\mathbf{T} : \mathbf{A} \to 2_*^{\mathbf{B}}$, where $\mathbf{A}, \mathbf{B}$ denote arbitrary alphabets and $2_*^{\mathbf{B}}$ is the power set of $\mathbf{B}$ without the empty set.
2) An *uncertain wiretap channel* is a pair $(\mathbf{T}_B, \mathbf{T}_C)$ of uncertain channels, where for finite alphabets $\mathbf{A}, \mathbf{B}, \mathbf{C}$, the channels are $\mathbf{T}_B : \mathbf{A} \to 2_*^{\mathbf{B}}$ and $\mathbf{T}_C : \mathbf{A} \to 2_*^{\mathbf{C}}$.

The reason why we only allow uncertain wiretap channels with finite alphabets is to exclude pathological cases. We define uncertain channels for general alphabets because we will eventually also define the quantizer as an uncertain channel, where the input alphabet obviously is infinite. An uncertain wiretap channel is shown in Fig. 2.

For every input $a \in \mathbf{A}$, the uncertain channel $\mathbf{T}$ generates a nonempty set $\mathbf{T}(a)$ of possible outputs. If a is an input into $\mathbf{T}$, then one element of $\mathbf{T}(a)$ will be the output. The outputs are not weighted according to any probabilities. Note that every mapping $g : \mathbf{A} \to \mathbf{B}$ can be regarded as an uncertain channel each of whose output sets is a singleton. We write

$$\operatorname{ran}(\mathbf{T}) := \bigcup_{a \in \mathbf{A}} \mathbf{T}(a)$$

for the set of outputs of $\mathbf{T}$ which can be generated by $\mathbf{T}$. Given two uncertain channels $\mathbf{F} : \mathbf{M} \to 2_*^{\mathbf{A}}$ and $\mathbf{T} : \mathbf{A} \to 2_*^{\mathbf{B}}$, we define the concatenation $\mathbf{T} \circ \mathbf{F} : \mathbf{M} \to 2_*^{\mathbf{B}}$ of $\mathbf{F}$ and $\mathbf{T}$ by the rule

$$(\mathbf{T} \circ \mathbf{F})(m) := \mathbf{T}(\mathbf{F}(m)) := \bigcup_{a \in \mathbf{F}(m)} \mathbf{T}(a),$$

for every $m \in \mathbf{M}$.

If a sequence $a(0:n)$ is to be transmitted over $n+1$ time slots, where a different channel $\mathbf{T}_i$ is used at each time $i \in \{0, \dots, n\}$, we define the set of possible outputs to be

$$\mathbf{T}_{0:n}(a(0:n)) := \mathbf{T}_0(a(0)) \times \cdots \times \mathbf{T}_n(a(n)). \tag{2}$$

$\mathbf{T}_{0:n} : \mathbf{A}^{n+1} \to 2_*^{\mathbf{B}^{n+1}}$ is an uncertain channel as well. If $\mathbf{T}_0 = \ldots = \mathbf{T}_n$, then we write $\mathbf{T}^{n+1} := \mathbf{T}_{0:n}$. Property (2) implies that channel inputs at different times do not influence each other. We also define the *reverse* $\mathbf{T}^{-1} : \mathbf{B} \to 2_*^{\mathbf{A}}$ of an uncertain channel $\mathbf{T} : \mathbf{A} \to 2_*^{\mathbf{B}}$ via

$$\mathbf{T}^{-1}(b) := \{a \in \mathbf{A} : b \in \mathbf{T}(a)\}.$$

Clearly, for channels $\mathbf{T}_0, \dots, \mathbf{T}_n$, the reverse $\mathbf{T}_{0:n}^{-1}$ of $\mathbf{T}_{0:n}$ satisfies

$$\mathbf{T}_{0:n}(b(0:n)) = \mathbf{T}_0^{-1}(b(0)) \times \cdots \times \mathbf{T}_n^{-1}(b(n)).$$

We denote the reverse of a channel $\mathbf{T}^n$ by $\mathbf{T}^{-n}$.

Now let $(\mathbf{T}_B, \mathbf{T}_C)$ be the uncertain wiretap channel which has to be employed by the encoder in order to send system state information to the estimator. The estimator, i.e., the receiver of $\mathbf{T}_B$ should be able to decode the transmitted information, whereas the eavesdropper which listens on channel $\mathbf{T}_C$ should obtain as little information as possible.

First we describe how $\mathbf{T}_B$ is used to convey information about the system states to the estimator. We assume that for each step of the evolution of the dynamical system, the channel can be used exactly once.

Definition 2: A *transmission scheme* is a sequence of quadruples $(\nu_k, n_k, f_k, \varphi_k)_{k=0}^{\infty}$, where $(\nu_k)_{k=0}^{\infty}$ and $(n_k)_{k=0}^{\infty}$ are bounded sequences of positive integers such that $n_k > \nu_k$ for at most finitely many k, and setting $\tau_k := \sum_{i=0}^{k} \nu_i$ and $t_k := \sum_{i=0}^{k} n_i$, for every $k \geq 0$

- the k-th encoder $f_k : \mathbb{R}^{\tau_k} \to 2_*^{\mathbf{A}^{n_k}}$ is an uncertain channel,
- the k-th decoder $\varphi_k : \mathbf{B}^{t_k} \to \mathbb{R}^{\nu_k}$ is an ordinary mapping.

f_k takes the system path until time τ_k as input and maps this into a codeword of length n_k. We allow f_k to be an uncertain channel because the optimal encoders for uncertain wiretap channels in general are uncertain channels as well, see after Definition 4. The decoder φ_k takes the first t_k outputs of $\mathbf{T}_B$ and calculates an estimate of the states $x(\tau_{k-1}), \dots, x(\tau_k - 1)$, which have not been estimated before. When we define the performance criterion for a transmission scheme, it will be seen that by not allowing φ_k to be an uncertain channel we do not lose generality.

Allowing the sequences $(\nu_k)_{k=0}^{\infty}$ and $(n_k)_{k=0}^{\infty}$ to differ gives us some additional flexibility in coding, but it is not crucial for the analysis. It can be proved that the requirements on the two sequences ensure that the delay D of the transmission scheme is finite, in the sense that the estimator waits D time slots and then outputs an estimate $\hat{x}(t)$ of $x(t)$ at time $t + D$ for all $t \geq 0$. However, in this paper we will not be concerned with the exact value of the delay and concentrate only on the reliability and security properties of a transmission scheme, as we will define next.

We now introduce the notation which is necessary to define the concepts of reliability and security. For a set $\mathcal{X} \subset \mathbb{R}$, we define $\operatorname{vol}(\mathcal{X})$ to be the Lebesgue measure of $\mathcal{X}$ and $\operatorname{diam}(\mathcal{X})$ to be its diameter, i.e., $\operatorname{diam}(\mathcal{X}) = \sup\{|x - x'| : x, x' \in \mathcal{X}\}$. For a set $\mathcal{A} \subset \mathbb{R}^{t+1}$ and $s \in \{0, \dots, t\}$, we set $\mathcal{A}|_s$ to be the set $\{a(s) \in \mathbb{R} : a(0:t) \in \mathcal{A}\}$. We also define

$$\operatorname{diam}(\mathcal{A}) := \max_{0 \leq s \leq t} \operatorname{diam}(\mathcal{A}|_s).$$

An output sequence $b(0:n-1) \in \operatorname{ran}(\mathbf{T}_B^n)$ and an output sequence $c(0:n-1) \in \operatorname{ran}(\mathbf{T}_C^n)$ are called *compatible* if there exists an $a(0:n-1) \in \mathbf{A}^n$ such that $b(0:n-1) \in \mathbf{T}_B^n(a(0:n-1))$ and $c(0:n-1) \in \mathbf{T}_C^n(a(0:n-1))$.

A transmission scheme only defines a decoder at the output of the estimator's channel $\mathbf{T}_B$. But every system path $x(0:\infty)$ also generates a sequence $c(0:\infty)$ of eavesdropper outputs. The security criterion will be such that transmitted information has to be secure no matter which decoding strategy the eavesdropper applies.

Definition 3:

1) A transmission scheme $(\nu_k, n_k, f_k, \varphi_k)_{k=0}^{\infty}$ is called *reliable* if there exists a constant $\varrho \in [0, \infty)$ such that for every $k \geq 0$ and every $b(0:t_k - 1) \in \mathbf{B}^{t_k}$,

$$\operatorname{diam}(f_{0:k}^{-1}(\mathbf{T}_B^{-t_k}(b(0:t_k-1)))) \leq \varrho, \tag{3}$$
$$\varphi_k(b(0:t_k-1)) \in f_{0:k}^{-1}(\mathbf{T}_B^{-t_k}(b(0:t_k-1))). \tag{4}$$

2) For $\sigma > 0$, a transmission scheme $(\nu_k, n_k, f_k, \varphi_k)_{k=0}^{\infty}$ is called σ-*weakly secure* if there exists a $\gamma > 0$ such that for sufficiently large k, all output sequences $b(0:t_k - 1) \in \operatorname{ran}(\mathbf{T}_B^{t_k} \circ f_{0:k})$ and every $c(0:t_k-1) \in \mathbf{C}^{t_k}$ which is compatible with $b(0:t_k-1)$,

$$\frac{\operatorname{vol}(f_{0:k}^{-1}(\mathbf{T}_C^{-t_k}(c(0:t_k-1)))|_{\tau_k - 1})}{\operatorname{vol}(f_{0:k}^{-1}(\mathbf{T}_B^{-t_k}(b(0:t_k-1)))|_{\tau_k-1})} \geq \gamma \cdot 2^{\sigma\tau_k}.$$

3) For $\sigma > 0$, a transmission scheme $(\nu_k, n_k, f_k, \varphi_k)_{k=0}^{\infty}$ is called σ-*strongly secure* if there exists a $\gamma > 0$ such that for sufficiently large k and for all eavesdropper output sequences $c(0:t_k-1) \in \operatorname{ran}(\mathbf{T}_C^{t_k} \circ f_{0:k})$,

$$\operatorname{vol}(f_{0:k}^{-1}(\mathbf{T}_C^{-t_k}(c(0:t_k-1)))|_{\tau_k-1}) \geq \gamma \cdot 2^{\sigma\tau_k}.$$

The reason why the definition of reliability contains two conditions is that what effectively has to be controlled is the diameter of the set of possible states, which is (3). Condition (4) mainly ensures that the estimation error does not exceed the diameter of this set. It shows that one would not gain anything by allowing φ_k to be an uncertain channel. In the definition of both weak and strong security, for every $k \geq 0$, a volume requirement is only made at the last time slot $\tau_k - 1$ of the corresponding encoding block. However, since the blocklength sequences $(\nu_k)_{k=0}^{\infty}$ and $(n_k)_{k=0}^{\infty}$ are bounded, these restrictions also extend to time instances t between the times of the form $\tau_k - 1$. Observe that, if a transmission scheme is reliable and strongly σ-secure for some $\sigma > 0$, then it is also weakly σ'-secure for some $\sigma' > 0$. This justifies the terms "weak" and "strong".

Our goal is to relate the existence of a transmission scheme which is both reliable and (weakly or strongly) secure to the secrecy capacity $C_0(\mathbf{T}_B, \mathbf{T}_C)$ of the uncertain wiretap channel $(\mathbf{T}_B, \mathbf{T}_C)$. The secrecy capacity was introduced in [13] and will be defined after some necessary preparations.

Definition 4: A *zero-error wiretap* (n, M)*-code for* $(\mathbf{T}_B, \mathbf{T}_C)$ is an uncertain channel $\mathbf{F} : \{0, \dots, M-1\} \to 2_*^{\mathbf{A}^n}$ satisfying

1) $\mathbf{T}_B^n(\mathbf{F}(m)) \cap \mathbf{T}_B^n(\mathbf{F}(m')) = \varnothing$ for all $m, m' \in \{0, \dots, M-1\}$ with $m \neq m'$,
2) for every $c(0:n-1) \in \operatorname{ran}(\mathbf{T}_C^n \circ \mathbf{F})$ there exist messages $m, m' \in \{0, \dots, M-1\}$ such that $m \neq m'$ and $c(0:n-1) \in \mathbf{T}_C^n(\mathbf{F}(m)) \cap \mathbf{T}_C^n(\mathbf{F}(m'))$.

Every m is called a *message* and n is called the *blocklength* of $\mathbf{F}$.

The first condition on $\mathbf{F}$ says that the receiver of $\mathbf{T}_B$ should without any ambiguity be able to recover the message which was sent: No possible channel output can be generated by more than one message. The second condition on $\mathbf{F}$ means that the exact opposite should hold for every output sequence $c(0:n-1)$ the eavesdropper might receive: It should be possible to generate this output sequence by at least two messages. Therefore, the eavesdropper cannot tell which message actually was sent. That $\mathbf{F}$ can be an uncertain channel instead of a normal mapping has to do with the fact that using uncertain channels can strictly increase the maximal number of messages for fixed n, see [13].

Definition 5: For $n \geq 1$, let $N(n)$ be the maximal M for which there exists a zero-error wiretap (n, M)-code for $(\mathbf{T}_B, \mathbf{T}_C)$. The number

$$C_0(\mathbf{T}_B, \mathbf{T}_C) := \sup_n \frac{\log N(n)}{n}$$

is then called the *zero-error secrecy capacity of* $(\mathbf{T}_B, \mathbf{T}_C)$.

By the superadditivity of the sequence $(\log N(n))_{n=1}^{\infty}$ and Fekete's lemma ([3], see also [12]), the supremum in the definition of $C_0(\mathbf{T}_B, \mathbf{T}_C)$ can be replaced by $\lim_{n\to\infty}$.

When zero-error wiretap codes are applied in the construction of secure transmission schemes, the achievable degree of security also depends on how many messages can possibly generate a given eavesdropper channel output.

Definition 6: An (n, M, ζ)-code is a zero-error wiretap (n, M)-code $\mathbf{F}$ where $|\mathrm{ran}(\mathbf{F}^{-1}(\mathbf{T}_C^{-n}(c(0:n-1))))| \geq \zeta$ for every $c(0:n-1) \in \mathrm{ran}(\mathbf{T}_C \circ \mathbf{F})$.

Obviously $\zeta \geq 2$. With an (n, M, ζ)-code, the eavesdropper has for each of its outputs at least ζ messages that could have generated this output. The interplay between M and ζ is crucial for our main results.

Definition 7: By $R_0(\mathbf{T}_B, \mathbf{T}_C)$ we denote the set of pairs (r, z) of nonnegative real numbers which satisfy that for every $\varepsilon > 0$ and sufficiently large n there exists an (n, M, ζ)-code with

$$\frac{\log M}{n} \geq r - \varepsilon, \quad \frac{\log \zeta}{n} \geq z - \varepsilon.$$

If we concatenate an (n, M, ζ)-code with itself to form a zero-error wiretap code of blocklength $2n$, note that this is a $(2n, M^2, \zeta^2)$-code. Therefore it makes sense to consider the exponential growth rate of the maximal ζ as n tends to infinity, which is positive if $C_0(\mathbf{T}_B, \mathbf{T}_C) > 0$. The set $R_0(\mathbf{T}_B, \mathbf{T}_C)$ actually contains those pairs of growth rates which can be achieved jointly by M and ζ. Its interior is nonempty, because to every positive $r < C_0(\mathbf{T}_B, \mathbf{T}_C)$ there exists a positive z with $(r, z) \in R_0(\mathbf{T}_B, \mathbf{T}_C)$.

III. MAIN RESULTS AND PROOFS

A. Main Results

Theorem 1: Assume that $\Omega = 0$ and $C_0(\mathbf{T}_B, \mathbf{T}_C) > \log \lambda$.

1) For every $(r, z) \in R_0(\mathbf{T}_B, \mathbf{T}_C)$ with $r > \log \lambda$ and $z > 0$ and every sufficiently small $\varepsilon > 0$ there exists a transmission scheme $(\nu_k, n_k, f_k, \varphi_k)_{k=0}^{\infty}$ which is reliable and $(z - \varepsilon)$-weakly secure.
2) For every $(r, z) \in R_0(\mathbf{T}_B, \mathbf{T}_C)$ with $r > \log \lambda$ and $z > r - \log \lambda$ and every sufficiently small $\varepsilon > 0$, there exists a transmission scheme $(\nu_k, n_k, f_k, \varphi_k)_{k=0}^{\infty}$ which is reliable and $(z - r + \log \lambda - \varepsilon)$-strongly secure.

Theorem 2: Assume that $\Omega > 0$ and $C_0(\mathbf{T}_B, \mathbf{T}_C) > \log \lambda$. For every $(r, z) \in R_0(\mathbf{T}_B, \mathbf{T}_C)$ with $r > \log \lambda$, define

$$\sigma := \frac{z \log \lambda}{r + 3 \log \lambda}.$$

Then for every $0 < \varepsilon < \sigma$ there exists a transmission scheme $(\nu_k, n_k, f_k, \varphi_k)_{k=0}^{\infty}$ which is reliable and $(\sigma - \varepsilon)$-strongly secure.

We do not quantify the estimation error ϱ in the definition of reliability, because it heavily depends on the channel and the delay. To obtain more insight into this question, a precise analysis of the function $N(n)$ would be necessary. However, this is currently out of reach for general channels even if one disregards the security condition 2) in Definition 4, cf. [6].

B. Blocklength-1 Analysis

In this subsection, we analyze the system under the assumption that a zero-error wiretap $(1, M, \zeta)$-code exists. This does not have to be the case even if $C_0(\mathbf{T}_B, \mathbf{T}_C) > 0$, as was shown in [13], but it allows us to concentrate on the properties of the dynamical system. In Subsection III-C, the results of this subsection will be applied to codes with larger blocklengths and the correspondingly subsampled dynamical system to obtain Theorems 1 and 2.

We assume the existence of a $(1, M, \zeta)$-code $\mathbf{F}$ (obviously with $M \geq \zeta \geq 2$). Once we have defined a quantizer, we are ready to derive the main technical results on which the proofs of Theorems 1 and 2 are based. We first define the recursive partition system underlying the quantizer. Define $[A(m(0:-1)), B(m(0:-1))] := I_0$ and $\hat{x}(m(0:-1))$ to be the mid point of I_0. Then for every $t \geq 0$ and every sequence $m(0:t) \in M^{t+1}$, we recursively set

$$P(m(0:t)) := A(m(0:t-1)) \tag{5}$$
$$+ \frac{B(m(0:t-1)) - A(m(0:t-1))}{M}[m(t), m(t)+1],$$
$$\hat{x}(m(0:t)) := \text{mid point of } P(m(0:t)), \tag{6}$$
$$[A(m(0:t)), B(m(0:t))] := \lambda P(m(0:t)) + \left[-\frac{\Omega}{2}, \frac{\Omega}{2}\right]. \tag{7}$$

Then we can recursively for $t \geq 0$ define the quantizer q, an uncertain channel, by

$$q(x(0)) := \{m(0) : x(0) \in P(m(0))\},$$
$$q(x(t), m(0:t-1)) := \{m(t) : x(t) \in P(m(0:t))\},$$

where $m(0:t-1)$ is the sequence of earlier quantizer outputs.

For every $t \geq 0$, the interval $[A(m(0:t-1)), B(m(0:t-1))]$ is the set of states $x(t)$ that are possible according to the quantizer index sequence $m(0:t-1)$ generated by the system until time $t-1$. In other words, it represents the estimator's knowledge about $x(t)$ after receiving the messages $m(0),\dots,m(t-1)$ and before receiving $m(t)$. After reception of $m(t)$, the estimator knows that $x(t) \in P(m(0:t))$ and thus sends a message $m(t) \in q(x(0:t))$ through the channel. Note that $|q(x(0:t))| = 2$ if $x(t)$ lies on the boundary between two neighboring *quantization intervals* $P(m(0:t-1),m)$ and $P(m(0:t-1),m+1)$; otherwise, $|q(x(0:t))| = 1$. Given $m(0:t)$, the estimate of $x(t)$ is $\hat{x}(m(0:t))$. Clearly, every path $x(0:\infty)$ generates an infinite sequence $m(0:\infty)$. Moreover, note that the path $\hat{x}(m(0:\infty))$ generates $m(0:\infty)$.

The following lemma from [13] takes care of the estimation error by bounding the length of the quantization intervals.

Lemma 1: For $t \geq 0$ and $m(0:t) \in \{0,\dots,M-1\}^{t+1}$,

$$|P(m(0:t))| = \frac{\lambda}{M}|P(m(0:t-1))| + \frac{\Omega}{M}. \tag{8}$$

In particular, if $M > \lambda$, then

$$|P(m(0:t))| = \frac{\Omega}{M-\lambda} + \left(\frac{\lambda}{M}\right)^t \left(\frac{|I_0|}{M} - \frac{\Omega}{M-\lambda}\right) \tag{9}$$

and

$$\sup_t |P(m(0:t))| = \max\left\{\frac{|I_0|}{M}, \frac{\Omega}{M-\lambda}\right\}. \tag{10}$$

Henceforth, we will frequently just write $|P(m(0:t))|$ without specifying which $m(0:t)$ is meant. This is justified by the fact that $|P(m(0:t))|$ only depends on t, as shown in Lemma 1. Security in the case $\Omega = 0$ uses the following lemma.

Lemma 2: If $\Omega = 0$, then at each time $t \geq 0$, the interiors of the intervals $P(m(0:t))$ are disjoint, where $m(0:t)$ ranges over $\{0,\dots,M-1\}^{t+1}$.

Proof: Omitted due to space constraints. ■

With the above quantizer q, the estimator will in each time step t receive one message $m(t)$ by transmission through $\mathbf{T}_B$ using the zero-error wiretap code $\mathbf{F}$ and know that the true state is contained in $P(m(0:t))$. In contrast to this, according to the eavesdropper's information, at each time step at least ζ messages could be the true one. Thus after $t+1$ steps, there are ζ^{t+1} possible message paths $m(0:t)$. As the interiors of the corresponding quantizer sets are disjoint according to Lemma 2, the volume of the set of possible states according to the eavesdropper's information therefore is at least $\zeta^{t+1}|P(m(0:t))|$. This will provide weak security and, if ζ is large enough, also strong security.

To obtain security in the case of $\Omega > 0$ is more complicated. This is due to the fact that the disjointness of the quantizer intervals as proved above for the case $\Omega = 0$ breaks down for $\Omega > 0$.

Example 1: Let $\lambda = 2$, $\Omega = 2$, $I_0 = [-1,1]$ and $M = 4$. Then $P(2) = [-\frac{1}{2}, 0]$ and $P(3) = [0, \frac{1}{2}]$. In the next step, one has

$$P(2,3) = \left[-\frac{1}{2}, \frac{1}{4}\right], \quad P(3,2) = \left[-\frac{1}{4}, \frac{1}{2}\right],$$

so $P(2,3)$ and $P(3,2)$ are not disjoint. The closer a state $x(t)$ is to the origin (and the larger t), the more paths there are which can be in this particular state at time t.

The above example shows that one can only hope to obtain disjoint quantizer sets for a subset of message sequences. Let $c(0:\infty)$ be an output sequence of the channel from encoder to eavesdropper. For every $t \geq 0$, the application of the $(1,M,\zeta)$-code $\mathbf{F}$ gives rise to a set $\mathcal{M}_t := \{m_{t,1} < m_{t,2} < \dots < m_{t,\zeta}\} \subset \mathbf{T}_C^{-1}(c(t))$ of ζ messages which could have generated the t-th output seen by the eavesdropper. Write $\Xi := \{1,\dots,\zeta\}$ and fix a $T \geq 1$. For $j \geq 1$ and $\xi(1:j) \in \Xi^j$, we define the message sequence $m_{\xi(1:j)}(0:jT-1)$ by

$$m_{\xi(1:j)}(s) = m_{\xi(i),s}$$

if $1 \leq i \leq j$ and $(i-1)T \leq s \leq iT-1$. All of these message sequences are candidates for being the true sequence according to the eavesdropper's information. In addition, on the j-th block of components $(j-1)T,\dots,jT-1$, the sequences with the same history $m_{\xi(1:j-1)}(0:(j-1)T-1)$ are a componentwise ordered set of ζ message sequences. The corresponding quantizer intervals $P(m_{\xi(1:j)}(0:jT-1))$ will therefore diverge due to the system instability. If T is chosen sufficiently large, this implies disjointness of the quantizer intervals.

Lemma 3: Let $\Omega > 0$ and assume $M > \lambda$. Choose a $T \in \mathbb{N}$ satisfying

$$T \geq 2 + \frac{2\log M}{\log\lambda} - \frac{\log(M-\lambda)}{\log\lambda}. \tag{11}$$

Then for every $j \geq 1$, the interiors of the sets $P(m_{\xi(1:j)}(0:jT-1))$, where $\xi(1:j)$ ranges over Ξ^j, are disjoint.

Proof: Omitted due to space restrictions. ■

C. Proof of the Theorems

We begin by describing the transmission schemes, which are the same for the cases $\Omega = 0$ and $\Omega > 0$. Let $(r,z) \in R_0(\mathbf{T}_B,\mathbf{T}_C)$ with $r > \log\lambda$ and let $\varepsilon < \min\{r - \log\lambda, z\}$. For sufficiently large n there is then a zero-error wiretap (n,M,ζ)-code $\mathbf{F}$ with

$$2^{nr} \geq M > 2^{n(r-\varepsilon)}, \quad \zeta > 2^{n(z-\varepsilon)}. \tag{12}$$

In particular, $M > \lambda^n$. We now consider the system

$$x^{(n)}(k+1) = \lambda^n x(k) + w^{(n)}(k), \quad x^{(n)}(0) \in I_0,$$

where the plant disturbances $w^{(n)}(k)$ are given by

$$w^{(n)}(k) = \sum_{i=0}^{n-1} \lambda^{n-j-1} w(kn+j)$$

and may assume any value in the interval $[-\Omega^{(n)}/2, \Omega^{(n)}/2]$ with

$$\Omega^{(n)} = \frac{\Omega}{\lambda-1}(\lambda^n - 1).$$

This is the *n-sampled version* of our original dynamical system. It satisfies $x(kn) = x^{(n)}(k)$ for all $k \geq 0$. It has the same form as our original system (1). Therefore, a quantizer $q^{(n)}$ can be defined for this system in the same way as for (1).

Note that $q^{(1)} = q$. Further, $q^{(n)}(x^{(n)}(0:k))$ is a function of $x(0:kn)$, but it only depends on the states at times $0, n, \dots, kn$.

Next we define a transmission scheme $(\nu_k, n_k, f_k, \varphi_k)_{k=0}^{\infty}$. We set $\nu_0 = 1$ and $\nu_k = n$ for all $k \geq 1$ as well as $n_k = n$ for all $k \geq 0$. Thus $\nu_k \neq n_k$ only for $k = 0$. The reason for this is that in the previous section III-B, the analysis naturally implied quantization of the initial state. This carries over to the quantizers $q^{(n)}$, which makes an observation blocklength of $\nu_0 = 1$ right at time $t = 0$ necessary.

The encoders f_k are the concatenation of the quantizer with $\mathbf{F}$. Given a quantizer output, we just encode this output using the zero-error wiretap channel $\mathbf{F}$. Hence we can write

$$f_k(x(0:kn)) = \mathbf{F}(q^{(n)}(x^{(n)}(0:k))).$$

At the estimator node, the decoder φ_k, which knows the history $m(0:k-1)$, decodes the k-th message $m(k)$ and maps it to $\hat{x}^{(n)}(m(0:k))$.

The encoded values are transmitted error-less to the estimator. According to Lemma 1, the diameter of the sets $P^{(n)}(m(0:k))$ is upper-bounded by

$$\max\left\{\frac{\Omega^{(n)}}{M-\lambda^n}, \frac{|I_0|}{M}\right\} = \max\left\{\frac{\Omega}{\lambda-1}\frac{\lambda^n-1}{M-\lambda^n}, \frac{|I_0|}{M}\right\}.$$

Thus our transmission scheme is reliable. Next we prove the security assertions.

Security for $\Omega = 0$*:* Hardly anything remains to be proved here. By Lemma 2, for given k, the interiors of all possible $P^{(n)}(m(0:k))$ are disjoint. Assume the true message sequence equals $m(0:k)$ and that the eavesdropper obtains the channel output sequence $c(0:t_k-1) = c(0:(k+1)n-1)$. For every $0 \leq i \leq k$ and every received output block $c(in:(i+1)n-1)$, since the encoder applies an (n, M, ζ)-code, there are at least ζ possible messages which could have generated this output sequence. Thus after time $t_k - 1$, the eavesdropper has at least ζ^{k+1} possible message sequences to choose from, and the state could be in one of ζ^{k+1} possible intervals of length $|P^{(n)}(m(0:k))|$. With $\tau_k = kn+1$, this provides $\log\zeta^{1/n}$-weak security, and hence with (12) also $(z-\varepsilon)$-weak security.

For strong security, observe that by Lemma 1

$$\zeta^{k+1}|P^{(n)}(m(0:k))| = \left(\frac{\zeta\lambda^n}{M}\right)^{k+1}\frac{|I_0|}{\lambda^n}.$$

This tends to infinity at exponential speed if $\zeta > M/\lambda^n$. Again observing that $\tau_k = kn+1$, σ-strong security follows with the σ claimed in the statement of Theorem 1.

Security for $\Omega > 0$*:* Define

$$T^{(n)} := 2 + \left\lceil \frac{2\log M}{\log\lambda^n} - \frac{\log(M-\lambda^n)}{\log\lambda^n} \right\rceil.$$

$T^{(n)}$ satisfies the requirements for Lemma 3 to be valid for the n-sampled system and quantizer $q^{(n)}$. For a given eavesdropper output sequence $c(0:jT^{(n)}n-1)$, there are $\zeta^{jT^{(n)}}$ possible message sequences which could have generated this output sequence. Out of these, the ζ^j message sequences of the form $m_{\xi(1:j)}(0:jT^{(n)}-1)$ produce quantizing sets $P^{(n)}(m_{\xi(1:j)}(0:jT^{(n)}-1))$ with disjoint interiors. This gives $\zeta^{1/(nT^{(n)})}$-strong security. Given an arbitrary $\delta > 0$, if n is large enough, then $\log(M-\lambda^n) \geq n(r-\delta)$. Thus for δ sufficiently small

$$T^{(n)} \leq 2 + \left\lceil \frac{r}{\log\lambda} + \delta \right\rceil \leq \frac{r}{\log\lambda} + 3.$$

Therefore we obtain from (12)

$$\zeta^{1/(nT^{(n)})} \geq \frac{z-\varepsilon}{T^{(n)}} \geq \frac{z\log\lambda}{r+3\log\lambda} - \varepsilon',$$

where ε' tends to 0 as δ and ε tend to zero. This shows strong security with the claimed parameter and concludes the proof of Theorem 2.

Remark 1: The above analyses require that the encoder has precise knowledge about λ. However, if the true λ is known to lie in a certain domain $[\lambda_-, \lambda_+]$, we expect all results to go through if $C_0(\mathbf{T}_B, \mathbf{T}_C) > \log\lambda_+$. Quantization would be done according to λ_+, but the divergence parameter σ would be determined by λ_-.

References

[1] C. De Persis and P. Tesi, "Input-to-state stabilizing control under denial-of-service," *IEEE Trans. Autom. Control*, vol. 60, no. 11, pp. 2930–2944, Nov 2015.

[2] 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, June 2014.

[3] M. Fekete, "Über die Verteilung der Wurzeln bei gewissen algebraischen Gleichungen mit ganzzahligen Koeffizienten," *Mathematische Zeitschrift*, vol. 17, no. 1, pp. 228–249, 1923.

[4] A. Gupta, A. Nayyar, C. Langbort, and T. Başar, "A dynamic transmitter-jammer game with asymmetric information," in *51st IEEE Conference on Decision and Control (CDC)*, pp. 6477–6482, Dec 2012.

[5] *IEEE Control Systems Magazine*, vol. 35, no. 1, February 2015, special issue on cyberphysical security in networked control systems.

[6] J. Körner and A. Orlitsky, "Zero-error information theory," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2207–2229, Oct 1998.

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

[8] A. S. Matveev and A. V. Savkin, "Shannon zero error capacity in the problems of state estimation and stabilization via noisy communication channels," *Int. Jour. Contr.*, vol. 80, no. 2, pp. 241–255, 2007.

[9] G. Nair, "A nonstochastic information theory for communication and state estimation," *IEEE Trans. Autom. Control*, vol. 58, no. 6, pp. 1497–1510, June 2013.

[10] M. Pajic, P. Tabuada, I. Lee, and G. J. Pappas, "Attack-resilient state estimation in the presence of noise," in *54th IEEE Conference on Decision and Control (CDC)*, pp. 5827–5832, Dec 2015.

[11] 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.

[12] J. H. van Lint and R. M. Wilson, *A Course in Combinatorics*. Cambridge University Press, 2001.

[13] M. Wiese, K. H. Johansson, T. J. Oechtering, P. Papadimitratos, H. Sandberg, and M. Skoglund, "Uncertain wiretap channels and secure estimation," in *Proc. 2016 International Symposium on Information Theory (ISIT '16)*, Barcelona, Spain, July 2016.