Uncertain Wiretap Channels and Secure Estimation

Moritz Wiese, Karl Henrik Johansson, Tobias J. Oechtering,
Panos Papadimitratos, Henrik Sandberg, Mikael Skoglund
ACCESS Linnaeus Centre, KTH Royal Institute of Technology, SE-10044 Stockholm, Sweden
{moritzw, kallej, oech, papadim, hsan, skoglund}@kth.se

***Abstract*—The zero-error secrecy capacity of uncertain wiretap channels is defined. If the sensor-estimator channel is perfect, it is also calculated. Further properties are discussed. The problem of estimating a dynamical system with nonstochastic disturbances is studied where the sensor is connected to the estimator and an eavesdropper via an uncertain wiretap channel. The estimator should obtain a uniformly bounded estimation error whereas the eavesdropper's error should tend to infinity. It is proved that the system can be estimated securely if the zero-error capacity of the sensor-estimator channel is strictly larger than the logarithm of the system's unstable pole and the zero-error secrecy capacity of the uncertain wiretap channel is positive.**

I. Introduction

If "independent noise" is assumed for every time step, it tends to be considered as stochastic in information theory. In contrast to this, robust control theory commonly treats dynamical systems with nonstochastic disturbances. In order to give a unified framework for the latter case, Nair has proposed a "nonstochastic information theory" [1]. The basic channel model in [1] is the newly introduced *uncertain channel*, a rule which determines which channel input can generate which channel outputs *without* weighting the possible outputs. For finite alphabets, every uncertain channel thus corresponds to a 0-1-matrix obtained from a stochastic matrix by replacing every positive entry by 1. Thus, uncertain channels are natural objects in zero-error information theory. Nair also introduced an analog to mutual information which plays the same role for the zero-error capacity of uncertain channels as mutual information for the capacity of discrete memoryless channels.

In [1], Nair applied his nonstochastic information theory to the problem of estimating an unstable scalar dynamical system with nonstochastic disturbances at a remote location which obtains sensor data through an uncertain channel $\mathbf{T}$. He showed that the estimation error can be bounded uniformly if the zero-error capacity $C_0(\mathbf{T})$ of $\mathbf{T}$ is strictly larger than the logarithm of the system's unstable pole $\lambda > 1$. This is "almost" necessary as well in the sense that $C_0(\mathbf{T}) \geq \log \lambda$ is required for uniform boundedness of the estimation error.

In this paper we add to the above problem an eavesdropper overhearing the communication between sensor and estimator via a second uncertain channel. For every eavesdropper output sequence there should be two system paths whose distance tends to infinity with increasing time. We call this the problem of secure estimation. A similar problem has been studied for the stochastic case in [2].

For our nonstochastic setting, this leads to the introduction of the uncertain wiretap channel: a pair $(\mathbf{T}_B, \mathbf{T}_C)$ of uncertain channels with common input alphabet. A zero-error wiretap code is a zero-error code for $\mathbf{T}_B$ such that every eavesdropper output word can be generated by at least two different messages. Surprisingly, positivity of the zero-error secrecy capacity $C_0(\mathbf{T}_B, \mathbf{T}_C)$ already is sufficient, in addition to Nair's sufficient condition for a bounded estimation error, in order for secure estimation to be possible.

The schemes for data transmission from the sensor to the estimator apply block codes. The error at inter-decoding times increases with communication delay. On the other hand, the estimation error at decoding times can be made to vanish asymptotically at the cost of increased delay. Similarly, we provide a lower bound on the speed of divergence for the eavesdropper's error which increases with increasing delay.

We calculate the secrecy capacity in the case of a perfect sensor-estimator channel. It either equals zero or the logarithm of the size of the input alphabet. An example shows that for general uncertain wiretap channels, no secure message transmission may be possible at blocklength 1, whereas a positive transmission rate is achieved for blocklengths ≥ 2. It also shows that encoders for zero-error wiretap codes in general have to be strictly uncertain channels, i. e. every message can be mapped to several possible codewords similar to the use of stochastic encoders for stochastic wiretap channels. We do not apply Nair's nonstochastic information-theoretic quantities in any of the analyses. Further, uncertain wiretap channels do not appear to provide new insights for the study of zero-error capacity, an overview of which is given in [3].

Section II describes the problems considered, Section III presents the results and Section IV contains a selection of proof sketches. Full proofs are available in the extended version [8] of this paper.

II. Model

A. Uncertain Channels

Let $\mathbf{A}, \mathbf{B}$ be finite alphabets. An *uncertain channel from* $\mathbf{A}$ *to* $\mathbf{B}$ is a mapping $\mathbf{T} : \mathbf{A} \to 2^{\mathbf{B}}_* := 2^{\mathbf{B}} \setminus \{\varnothing\}$. For any $a \in \mathbf{A}$, the set $\mathbf{T}(a)$ is the family of possible output values of the channel given the input a. Only one of the elements of $\mathbf{T}(a)$ will actually be attained when transmitting a. That $\mathbf{T}(a) \neq \varnothing$ for all a means that every input generates an output. We will write $\mathrm{ran}(\mathbf{T})$ for the set of possible outputs of $\mathbf{T}$, i. e. $\mathrm{ran}(\mathbf{T}) = \cup_{a \in \mathbf{A}} \mathbf{T}(a)$.

An *M-code* is a collection $\{\mathbf{F}(m) : 1 \leq m \leq M\}$ of nonempty and mutually disjoint subsets of $\mathbf{A}$. This is equivalent to an uncertain channel $\mathbf{F} : \{1, \ldots, M\} \to 2^{\mathbf{A}}_*$

978-1-5090-1806-2/16/$31.00 ©2016 IEEE

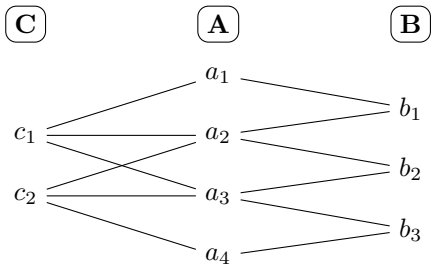

Fig. 1. An uncertain wiretap channel $(\mathbf{T}_B, \mathbf{T}_C)$. A line between a_i and b_j indicates that $b_j \in \mathbf{T}_B(a_i)$, similar for a_i and c_j. See Example 1 for an analysis of this particular channel.

with disjoint output sets, so we will often denote such a code by $\mathbf{F}$. The necessity of codes with $|\mathbf{F}(m)| \geq 2$ for some m is shown in Example 1. It is similar to the necessity of stochastic encoders for stochastic wiretap channels.

Obviously, first applying $\mathbf{F}$ and then $\mathbf{T}$ leads to a new uncertain channel $\mathbf{T} \circ \mathbf{F} : \{1, \dots, M\} \to 2^{\mathbf{B}}_*$ called the *composition of* $\mathbf{F}$ *and* $\mathbf{T}$. Formally, we have for any $m \in \{1, \dots, M\}$

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

A nonstochastic M-code $\mathbf{F}$ is called a *zero-error M-code for* $\mathbf{T}$ if for any $m, m' \in \{1, \dots, M\}$ with $m \neq m'$

$$\mathbf{T}(\mathbf{F}(m)) \cap \mathbf{T}(\mathbf{F}(m')) = \varnothing. \tag{1}$$

Thus every possible channel output $y \in \operatorname{ran}(\mathbf{T} \circ \mathbf{F})$ can be associated to a unique message m. For this to hold it is necessary that the sets $\mathbf{F}(m)$ be disjoint, which is the reason for this assumption in the definition of M-codes.

Given an additional finite alphabet $\mathbf{C}$, an *uncertain wiretap channel* is a pair of uncertain channels $(\mathbf{T}_B : \mathbf{A} \to 2^{\mathbf{B}}_*, \mathbf{T}_C : \mathbf{A} \to 2^{\mathbf{C}}_*)$. The interpretation is that the outputs of channel $\mathbf{T}_B$ are received by the intended receiver, whereas the outputs of $\mathbf{T}_C$ are heard by an eavesdropper. See Fig. 1 for an example.

An M-code $\mathbf{F}$ is called a *zero-error wiretap M-code for* $(\mathbf{T}_B, \mathbf{T}_C)$ if it is a zero-error code for $\mathbf{T}_B$ and additionally for every $c \in \operatorname{ran}(\mathbf{T}_C \circ \mathbf{F})$ there are messages $m \neq m'$ with

$$c \in \mathbf{T}_C(\mathbf{F}(m)) \cap \mathbf{T}_C(\mathbf{F}(m')). \tag{2}$$

Thus every output $c \in \operatorname{ran}(\mathbf{T}_C \circ \mathbf{F})$ can be generated by at least two possible messages.

We define the *n-fold product of an uncertain channel* $\mathbf{T} : \mathbf{A} \to 2^{\mathbf{B}}_*$ as the uncertain channel $\mathbf{T}^n : \mathbf{A}^n \to (2^{\mathbf{B}}_*)^n$,

$$\mathbf{T}^n(a^n) = \mathbf{T}(a_1) \times \cdots \times \mathbf{T}(a_n). \tag{3}$$

We call an M-code $\mathbf{F}$ on the alphabet $\mathbf{A}^n$ a *zero-error (M, n)-code for* $\mathbf{T}_B$ if (1) is satisfied with $\mathbf{T} \circ \mathbf{F}$ replaced by $\mathbf{T}^n_B \circ \mathbf{F}$. We call $\mathbf{F}$ a *zero-error wiretap (M, n)-code for* $(\mathbf{T}_B, \mathbf{T}_C)$ if $\mathbf{F}$ is a zero-error (M, n)-code for $\mathbf{T}_B$ and in addition, (2) is satisfied with $\mathbf{T}_C \circ \mathbf{F}$ replaced by $\mathbf{T}^n_C \circ \mathbf{F}$. We set $N_{\mathbf{T}_B}(n)$ to be the maximal M such that there exists a zero-error (M, n)-code for $\mathbf{T}_B$. Similarly, we set $N_{(\mathbf{T}_B, \mathbf{T}_C)}(n)$ to be the maximal M such that there exists a zero-error wiretap (M, n)-code for $(\mathbf{T}_B, \mathbf{T}_C)$. We then define

$$C_0(\mathbf{T}_B) = \sup_n \frac{\log N_{\mathbf{T}_B}(n)}{n}, \tag{4}$$

$$C_0(\mathbf{T}_B, \mathbf{T}_C) = \sup_n \frac{\log N_{(\mathbf{T}_B, \mathbf{T}_C)}(n)}{n}, \tag{5}$$

which are called the *zero-error capacity of* $\mathbf{T}_B$ and the *zero-error secrecy capacity of* $(\mathbf{T}_B, \mathbf{T}_C)$, respectively. Due to the superadditivity of the sequences $\log N_{\mathbf{T}_B}(n)$ and $\log N_{(\mathbf{T}_B, \mathbf{T}_C)}(n)$, the suprema in (4) and (5) can be replaced by limits by the well-known Fekete's lemma [4], see also [5].

B. The Unstable Dynamical System

Let $\lambda > 1$ and consider the real-valued system

$$x(t+1) = \lambda x(t) + w(t), \tag{6a}$$

$$x(0) = 0. \tag{6b}$$

where $w(t)$ is a nonstochastic disturbance with range $[-\Omega/2, \Omega/2]$ for some $\Omega > 0$. With

$$\tilde{\Omega}^{(t)} := \frac{\Omega}{\lambda - 1}(\lambda^t - 1), \tag{7}$$

the range of possible values of this system at time t equals $[-\tilde{\Omega}^{(t)}/2, \tilde{\Omega}^{(t)}/2]$, whose diameter grows exponentially in t. A sensor performs perfect state measurements, encodes them and sends them through an uncertain wiretap channel $(\mathbf{T}_B, \mathbf{T}_C)$. The dynamic system and the channel are synchronous, i. e. one symbol can be transmitted through the channel at every system time step. The goal is that the receiver of $\mathbf{T}_B$ (the estimator) be able to estimate the state with bounded estimation error and the eavesdropper's estimation error tend to infinity.

Formally, a *transmission scheme* $(n_k, f_k, \varphi_k)_{k=1}^{\infty}$ consists of a bounded sequence of positive natural numbers $(n_k)_{k=1}^{\infty}$ and, defining $t_k := \sum_{i=1}^{k} n_i$, for each $k \in \mathbb{N}$ an uncertain channel $f_k : \mathbb{R}^{t_k} \to 2^{\mathbf{X}^{n_k}}_*$ and a mapping $\varphi_k : \mathbf{Y}^{t_k} \to \mathbb{R}^{n_k}$. Every uncertain channel f_k maps the observations of the system state up till time t_k into one of several possible codewords of length n_k. The receiver of $\mathbf{T}_B$ uses φ_k to produce from all symbols received so far an estimate $\hat{x}(t_k), \dots, \hat{x}(t_{k+1} - 1)$ of the system states $x(t_k), \dots, x(t_{k+1} - 1)$.

The minimal delay which has to be tolerated is $\max_k n_k$. At this delay, the receiver has good estimates for the states at times t_k but has to extrapolate for $x(t_k + 1), \dots, x(t_{k+1} - 1)$. In particular, for the first $t_1 - 1$ steps of the evolution, the estimator has to rely on a rule which is independent of any observations and which we assume to estimate $\hat{x}(t) = 0$ ($0 \leq t \leq t_1 - 1$). Further, every system path $(x(t))_{t=0}^{\infty}$ generates a sequence $(c_t)_{t=1}^{\infty}$ of eavesdropper outputs.

Given a transmission scheme $(n_k, f_k, \varphi_k)_{k=1}^{\infty}$ and a sequence of estimates $(\hat{x}(t))_{t=0}^{\infty}$, we denote by $\mathbf{R}_B((\hat{x}(t))_{t=0}^{\infty})$ the set of system paths $(x(t))_{t=0}^{\infty}$ which using the transmission scheme can generate $(\hat{x}(t))_{t=0}^{\infty}$. One can consider $\mathbf{R}_B$ as an uncertain channel in the reverse direction with $\mathbb{R}^{\infty}$ as input and output alphabet. Similarly, for any infinite sequence $(c_t)_{t=1}^{\infty} \in \mathbf{Z}^{\infty}$ of eavesdropper outputs, we denote

by $\mathbf{R}_C((c_t)_{t=1}^{\infty})$ the set of system paths $(x(t))_{t=0}^{\infty}$ which can give rise to $(c_t)_{t=1}^{\infty}$.

For two sequences $(a_t)_{t=1}^{\infty}$ and $(b_t)_{t=1}^{\infty}$ let us define their distance to be $\|(a_t)-(b_t)\|_{\infty} := \sup_t |a_t - b_t|$. For a set S of sequences we define its diameter by

$$\operatorname{diam}(S) := \sup\{\|(a_t)-(b_t)\|_{\infty} : (a_t),(b_t) \in S\}.$$

The transmission scheme $(n_k, f_k, \varphi_k)_{k=1}^{\infty}$ is called *reliable* if the estimation error is bounded uniformly in the estimates, i. e. there exists a constant $\kappa > 0$ such that for every possible estimate sequence $(\hat{x}(t))_{t=0}^{\infty}$,

$$\sup\{\|(x(t))-(\hat{x}(t))\|_{\infty} : (x(t))_{t=0}^{\infty} \in \mathbf{R}_B((\hat{x}(t))_{t=0}^{\infty})\} \leq \kappa.$$

Further, $(n_k, f_k, \varphi_k)_{k=1}^{\infty}$ is called *secure* if for every sequence $(c_t)_{t=1}^{\infty} \subset \mathbf{C}^{\infty}$

$$\operatorname{diam}(\mathbf{R}_C((c_t)_{t=1}^{\infty})) = \infty.$$

Note that security is an asymptotic property due to the boundedness of the range of possible system states in any finite time horizon, cf. (7).

III. Results

A. Main Results

Theorem 1. *A reliable and secure transmission scheme exists if $C_0(\mathbf{T}_B) > \log\lambda$ and $C_0(\mathbf{T}_B, \mathbf{T}_C) > 0$.*

The main idea behind Theorem 1 is that the system's instability helps to achieve the goal of security as soon as a sufficiently large error on the eavesdropper side has been introduced at the beginning of transmission.

To apply Theorem 1, $C_0(\mathbf{T}_B)$ and $C_0(\mathbf{T}_B, \mathbf{T}_C)$ have to be known. However, the zero-error capacity $C_0(\mathbf{T}_B)$ is unknown for most channels except a few special cases, cf. [3]. Neither do we provide a general formula for $C_0(\mathbf{T}_B, \mathbf{T}_C)$ here. A solution can be given, though, when the calculation of $C_0(\mathbf{T}_B)$ is trivial.

Theorem 2. *If $\mathbf{T}_B$ is an injective function from $\mathbf{A}$ to $\mathbf{B}$, then $C_0(\mathbf{T}_B, \mathbf{T}_C) \in \{0, \log|\mathbf{A}|\}$. Further, $C_0(\mathbf{T}_B, \mathbf{T}_C) = 0$ if and only if there is no zero-error wiretap $(M,1)$-code for $(\mathbf{T}_B, \mathbf{T}_C)$ for any $M \geq 2$.*

For the proof of Theorem 2, it is sufficient to consider codes with $|\mathbf{F}(m)| = 1$ for all $1 \leq m \leq M$. The number of those elements of $\mathbf{A}^n$ which cannot be used as codewords grows exponentially, at a rate which is less than $\log|\mathbf{A}|$ if and only if there is no zero-error wiretap $(M,1)$-code for $(\mathbf{T}_B, \mathbf{T}_C)$ for any $M \geq 2$. Thus the number of elements of $\mathbf{A}^n$ that *can* be used either asymptotically grows with rate $\log|\mathbf{A}|$ or equals 0.

B. Estimation Error and Divergence Coefficient

As mentioned above, using a transmission scheme $(n_k, f_k, \varphi_k)_{k=1}^{\infty}$ with delay $\max_k n_k$, the estimates of system states $x(t)$ with $t \neq t_k$ $(k \in \mathbb{N})$ have to be extrapolated from the last good estimate. Thus the estimation error after a decoding time t_k grows exponentially until the next decoding time t_{k+1}. However, for any $\varepsilon > 0$ the estimation error at times $(t_k)_{k=1}^{\infty}$ can be made smaller than ε at least for large k if the inter-decoding intervals n_k $(k \in \mathbb{N})$ (and thus the inter-decoding estimate errors) are sufficiently large:

Lemma 1. *For every $\varepsilon > 0$ there exists a transmission scheme such that for every sequence $(\hat{x}(t))_{t=0}^{\infty}$ of estimates and every $(x(t))_{t=0}^{\infty} \in \mathbf{R}_B((\hat{x}(t))_{t=0}^{\infty})$,*

$$\limsup_{k\to\infty} |x(t_k) - \hat{x}(t_k)| \leq \varepsilon.$$

If $C_0(\mathbf{T}_B, \mathbf{T}_C) > \log\lambda$, then the limit superior can even be replaced by a supremum.

Another parameter of interest is the speed of divergence of the diameter of the set of possible system states given eavesdropper outputs $(c_t)_{t=1}^{T}$ as $T \to \infty$. Given a zero-error wiretap (M,n)-code $\mathbf{F}$, we define for every possible eavesdropper channel output $(c_t)_{t=1}^{n} \in \operatorname{ran}(\mathbf{T}_C^n \circ \mathbf{F})$

$$\begin{aligned}&\delta((c_t)_{t=1}^{n})\\ &= \max\{|m-m'|+1 : (c_t)_{t=1}^{n} \in \mathbf{T}_C^n(\mathbf{F}(m)) \cap \mathbf{T}_C^n(\mathbf{F}(m'))\}.\end{aligned}$$

Clearly $2 \leq \delta((c_t)_{t=1}^{n}) \leq M$. We then set

$$L := \min\{\delta((c_t)_{t=1}^{n}) : (c_t)_{t=1}^{n} \in \operatorname{ran}(\mathbf{T}_C^n \circ \mathbf{F})\}$$

and call $\mathbf{F}$ a (M, L, n)-code. We also define

$$\Delta_{(\mathbf{T}_B,\mathbf{T}_C)}(n) := \max\left\{\frac{L-1}{M-1} : \mathbf{F} \text{ is } (M,L,n)\text{-code}\right\}.$$

Clearly, $0 < \Delta_{(\mathbf{T}_B,\mathbf{T}_C)}(n) \leq 1$.

Lemma 2. *For every $\varepsilon > 0$ there exists a transmission scheme $(n_k, f_k, \varphi_k)_{k=1}^{\infty}$ such that for every eavesdropper output sequence $(c_t)_{t=1}^{\infty}$ there exist system paths $(x(t))_{t=1}^{\infty}, (x'(t))_{t=1}^{\infty} \in \mathbf{R}_C((c_t)_{t=1}^{\infty})$ satisfying*

$$\liminf_{T\to\infty} \frac{\|(x(t))_{t=1}^{T} - (x'(t))_{t=1}^{T}\|_{\infty}}{\lambda^T} \geq \frac{\Omega}{\lambda - 1} \sup_n \Delta_{(\mathbf{T}_B,\mathbf{T}_C)}(n) - \varepsilon.$$

The term on the right-hand side of the inequality in Lemma 2 is positive if ε is chosen small enough. The case $\sup_n \Delta_{(\mathbf{T}_B,\mathbf{T}_C)}(n) = 1$ corresponds to complete eavesdropper ignorance, cf. (7).

C. Uncertain Wiretap Channels

We first note that the divergence coefficient increases with increasing blocklength. Thus we find a trade-off between the growth rate for the eavesdropper's estimation error and delay:

Lemma 3. *If $C_0(\mathbf{T}_B, \mathbf{T}_C) > 0$, then*

$$\sup_n \Delta_{(\mathbf{T}_B,\mathbf{T}_C)}(n) = \lim_{n\to\infty} \Delta_{(\mathbf{T}_B,\mathbf{T}_C)}(n) > 0.$$

Next we have a closer look at the zero-error secrecy capacity of uncertain wiretap channels. To study the zero-error capacity of an uncertain channel $\mathbf{T} : \mathbf{A} \to 2_*^{\mathbf{B}}$, one associates to it the following graph $G(\mathbf{T})$: its vertex set equals $\mathbf{A}$ and an edge is drawn between $a, a' \in \mathbf{A}$ if $\mathbf{T}(a) \cap \mathbf{T}(a') \neq \varnothing$. In that case we write $a \sim a'$.

The graph $G(\mathbf{T}^n)$ corresponding to the n-fold product channel $\mathbf{T}^n$ (see (3)) is the *strong n-fold product of $G(\mathbf{T})$*

denoted by $G(\mathbf{T})^n$, in particular $G(\mathbf{T}^n) = G(\mathbf{T})^n$. Here for any graph G with vertex set $\mathbf{A}$, the strong product G^2 of G with itself is defined as follows: The vertex set of G^2 is $\mathbf{A}^2$ and $(a_1, a_2) \sim (a_1', a_2')$ if 1) $a_1 \sim a_1'$ and $a_2 = a_2'$ or 2) $a_2 \sim a_2'$ and $a_1 = a_1'$ or 3) $a_1 \sim a_1'$ and $a_2 \sim a_2'$.

Finding the zero-error capacity of $\mathbf{T}$ now amounts to finding the asymptotic behavior as $n \to \infty$ of the sizes of maximal independent systems of the graphs $G(\mathbf{T}^n)$, cf. [3]. We define an *independent system* in a graph as a set $\{\mathbf{F}(1), \dots, \mathbf{F}(M)\}$ of mutually disjoint subsets of the vertex set $\mathbf{A}$ such that no two vertices a, a' belonging to different subsets $\mathbf{F}(m) \neq \mathbf{F}(m')$ are connected by an edge.

To treat uncertain wiretap channels $(\mathbf{T}_B, \mathbf{T}_C)$, we consider a hypergraph structure $H(\mathbf{T}_C^n)$ induced on $\mathbf{A}^n$ in addition to the graph structure $G(\mathbf{T}_B^n)$. A hypergraph consists of a vertex set together with a set of subsets, called *hyperedges*, of this vertex set. The vertex set of $H(\mathbf{T}_C^n)$ equals $\mathbf{A}^n$. Every hyperedge is generated by a $(c_t)_{t=1}^n \in \mathbf{C}^n$: we set $e((c_t)_{t=1}^n) := \{(a_t)_{t=1}^n \in \mathbf{A}^n : (c_t)_{t=1}^n \in \mathbf{T}_C^n((a_t)_{t=1}^n)\}$.

It is easy to see that $H(\mathbf{T}_C^n)$ is the *n-fold square product* $H(\mathbf{T}_C)^n$, cf. [6]. For any hypergraph H with vertex set $\mathbf{A}$ and hyperedge set $\mathcal{E} \subset 2^{\mathbf{A}}$, the square product H^2 of H with itself is defined as follows: The vertex set of H^2 is $\mathbf{A}^2$ and the hyperedge set equals $\mathcal{E}^2 := \{e \times e' : e, e' \in \mathcal{E}\}$.

A zero-error wiretap (M, n)-code $\mathbf{F}$ then is nothing but a collection of disjoint subsets $\{\mathbf{F}(1), \dots, \mathbf{F}(M)\}$ of $\mathbf{A}^n$ satisfying the two following properties:

1) It is an independent system for $G(\mathbf{T}_B^n)$;
2) For every hyperedge e of $H(\mathbf{T}_C^n)$ there exist at least two different m, m' such that e has nonempty intersection with both $\mathbf{F}(m)$ and $\mathbf{F}(m')$.

This (hyper-)graph theoretic language is applied in the proof of Theorem 2. The following very interesting example gives additional insight into the nature of general uncertain wiretap channels and their secrecy capacity.

Example 1. Consider the wiretap channel $(\mathbf{T}_B, \mathbf{T}_C)$ from Fig. 1. $\mathbf{A}$ with $G(\mathbf{T}_B)$ and $H(\mathbf{T}_C)$ is depicted on the left of Fig. 2, $\mathbf{A}^2$ with $G(\mathbf{T}_B^2)$ and $H(\mathbf{T}_C^2)$ on its right. It is easy to check that there is no zero-error wiretap $(M, 1)$-code for any $M \geq 2$. On the other hand, a zero-error wiretap $(4, 2)$-code exists by choosing the codeword sets as indicated in Fig. 2. Therefore in the general case, in contrast to the situation in Lemma 2, there is no easy criterion an uncertain wiretap channel satisfies at blocklength 1 if and only if its zero-error secrecy capacity is positive.

This behavior of zero-error wiretap codes for general uncertain wiretap channels is remarkable when it is compared to the behavior of zero-error codes for uncertain channels: An uncertain channel $\mathbf{T}$ has $C_0(\mathbf{T}) > 0$ if and only if there exists an independent system for $G(\mathbf{T})$ with size ≥ 2. Similarly, a stochastic DMC has positive capacity if and only if its blocklength-1 transmission matrix does not have identical rows. For the secrecy capacity of stochastic wiretap channels, there is van Dijk's criterion [7] for positivity which concerns the blocklength-1 wiretap channel matrix and requires to check a certain function for concavity.

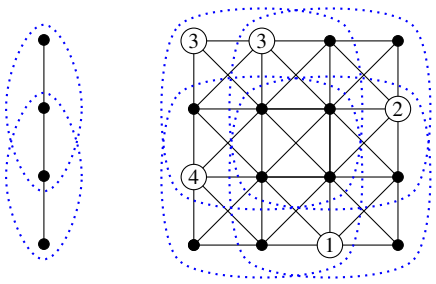

Fig. 2. $(\mathbf{T}_B, \mathbf{T}_C)$ from Example 1. Left: $\mathbf{A}$ with $G(\mathbf{T}_B)$ and $H(\mathbf{T}_C)$. Right: $\mathbf{A}^2$ with $G(\mathbf{T}_B^2)$ and $H(\mathbf{T}_C^2)$. Vertices connected by a solid black line are connected in $G(\mathbf{T}_B)$ or $G(\mathbf{T}_B^2)$, respectively. Vertices within the boundary of a blue dotted line belong to the same hyperedge of $H(\mathbf{T}_C)$ or $H(\mathbf{T}_C^2)$, respectively.

Observe also that in order to obtain a $(4, 2)$-code for the above channel, one message m has to be encoded into a set with $|\mathbf{F}(m)| \geq 2$. A simpler example illustrating the necessity of codes with $|\mathbf{F}(m)| \geq 2$ for some m is Example 2 in [8].

IV. Selected Proof Sketches

A. Proof of Theorem 1: Preliminaries

The first choice to make is the quantizer used by the sensor. For sufficient generality, we assume the rule (6a), but $x(0) \in I_0$ for some real interval I_0. Let $M \geq 2 \in \mathbb{N}$. Setting $P_{m_0,0} := I_0$, we recursively define for $t \geq 1$ and $1 \leq m \leq M$

$$[A(t), B(t)] = \lambda P_{m_{t-1},t-1} + \left[-\frac{\Omega}{2}, \frac{\Omega}{2}\right], \tag{8}$$

$$P_{m,t} = A(t) + (B(t) - A(t)) \left[\frac{m-1}{M}, \frac{m}{M}\right], \tag{9}$$

$$m_t = m \qquad \text{if } x(t) \in P_{m,t}. \tag{10}$$

In the definition of m_t, an uncertain mapping is applied to associate $x(t)$ to one of the two possible values if it lies on the boundary between two partition intervals $P_{m,t}, P_{m+1,t}$.

For every $t \in \mathbb{N}$, the interval $P_{m_t,t}$ is the set of system states which are possible at time t according to the sequence $(m_i)_{i=1}^t$. The interval $[A(t+1), B(t+1)]$ is the set of states the system could be in at time $t+1$ given that its state at time t is contained in $P_{m_t,t}$. The sets $P_{m,t+1} : 1 \leq m \leq M$ form an equal-sized partition of $[A(t+1), B(t+1)]$, and m_{t+1} is the index of the partition atom containing $x(t+1)$. Clearly, every path $(x(t))_{t=0}^\infty$ generates an infinite sequence $(m_t)_{t=1}^\infty$.

Lemma 4. *For every $t \in \mathbb{N}$ and $1 \leq m \leq M$,*

$$|P_{m,t}| = \left(\frac{\lambda}{M}\right)^t \left(|I_0| - \frac{\Omega}{M-\lambda}\right) + \frac{\Omega}{M-\lambda}. \tag{11}$$

Hence $\sup_t |P_{m_t,t}| < \infty$ if and only if $\lambda < M$. In that case

$$\sup_t |P_{m_t,t}| = \max\left\{|I_0|, \frac{\Omega}{M-\lambda}\right\}.$$

Lemma 4 is needed in the analysis of the intended receiver's estimation error and proved by induction over the recursion (8)-(10).

Next assume that we have two systems obeying (6a). The paths of one of them start in an interval I_0 and those of the other in an I_0' with $|I_0| = |I_0'|$. The same quantizer rules (8)-(10) are applied for both systems, generating sequences $(m_t)_{t=1}^{\infty}$ and $(m_t')_{t=1}^{\infty}$, respectively. For $t \geq 0$, denote by $\hat{x}(t)$ the mid point of $P_{m_t,t}$ and by $\hat{x}'(t)$ that of $P_{m_t',t}$. The following two lemmas are used in the security analysis of the scheme we are going to define and are proved using a recursion formula [8] for the sequences $(\hat{x}(t))_{t=0}^{\infty}, (\hat{x}'(t))_{t=0}^{\infty}$.

Lemma 5. *Let $L, M \geq 2$ and for every t let $1 \leq m_t' < m_t \leq M$ with $m_t - m_t' \geq L-1$. Then*

$$\liminf_{t\to\infty} \frac{\hat{x}(t) - \hat{x}'(t)}{\lambda^t} \geq \hat{x}(0) - \hat{x}'(0) + \frac{L-1}{M-1}\left(\frac{\Omega}{\lambda-1} + |I_0|\right).$$

Lemma 6. *Let $M \in \mathbb{N}$ and for every t let $1 \leq m_t, m_t' \leq M$. If*

$$|\hat{x}(0) - \hat{x}'(0)| > \frac{\Omega}{\lambda-1} + |I_0|, \tag{12}$$

then for every $t = 1, 2, \ldots$

$$\liminf_{t\to\infty} \frac{|\hat{x}(t) - \hat{x}'(t)|}{\lambda^t} \geq |\hat{x}(0) - \hat{x}'(0)| - \frac{\Omega}{\lambda-1} - |I_0|.$$

B. Proof of Theorem 1: Transmission Scheme

For any $n \geq 1$, let us introduce the *n-sampled system*

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

where $w^{(n)}(k)$ is a nonstochastic disturbance in the range $[-\tilde{\Omega}^{(n)}/2, \tilde{\Omega}^{(n)}/2]$ (cf. (7)). The n-sampled system describes the system (6) at the points $0, n, 2n, \ldots$

We restrict the presentation here to the more difficult case $C_0(\mathbf{T}_B, \mathbf{T}_C) \leq \log\lambda$. Choose n_1, n_2, M_1 such that $2 \leq M_1 < \max\{\lambda^{n_1}, N_{(\mathbf{T}_B,\mathbf{T}_C)}(n_1)\}$ and $N_{\mathbf{T}_B}(n_2) > \lambda^{n_2}$. Let $L \geq 2$ be chosen such that there exists a zero-error wiretap (M_1, L, n_1)-code $\mathbf{F}$ and let $\mathbf{G}$ be a zero-error $(N_{\mathbf{T}_B}(n_2), n_2)$-code.

We define a transmission scheme as follows: Do the construction (8)-(10) for the n_1-sampled system with M replaced by M_1 and $I_0 = \{0\}$, thus obtaining $A^{(n_1)}(k)$, $B^{(n_1)}(k)$, $P_{m,k}^{(n_1)}$, m_k (omitting the superscript (n_1) at m_k). For some $K \in \mathbb{N}$ to be chosen later and all $1 \leq k \leq K$, set

$$f_k(x(0), \ldots, x(kn_1)) = \mathbf{F}(m_k)$$

The intended receiver uses the mid point $\hat{x}(kn_1)$ of $P_{m_k,k}^{(n_1)}$ as estimate of $x(kn_1)$. For $k > K$, define $A^{(n_2)}(k-K)$, $B^{(n_2)}(k-K)$, $P_{m,k-K}^{(n_2)}$, m_{k-K} as in (8)-(10) with $I_0 = P_{m_K,K}^{(n_1)}$ (omitting the superscript (n_2) for m_{k-K}). Then set

$$f_k(x(0), \ldots, x(Kn_1 + (k-K)n_2)) = \mathbf{G}(m_{k-K}).$$

Decoding/estimating goes as in the first K steps.

As $N_{\mathbf{T}_B}(n_2) > \lambda^{n_2}$ it is clear by Lemma 4 that the estimation error for the intended receiver is bounded, so the transmission scheme is reliable. To prove its security, assume the eavesdropper receives a channel output sequence $(c_t)_{t=1}^{\infty}$. Lemma 5 implies the existence of paths $(x(t))_{t=0}^{\infty}, (x'(t))_{t=0}^{\infty}$ such that for any $\varepsilon > 0$ and sufficiently large K, the estimates at time Kn_1 have distance at least

$$\lambda^{Kn_1}\left(\frac{L-1}{M_1-1}\frac{\Omega}{\lambda-1} - \varepsilon\right). \tag{13}$$

By choosing K even larger if necessary, (12) is satisfied with its left-hand side replaced by (13) and on the right-hand side inserting $\tilde{\Omega}^{(n_2)}$ for Ω and $P_{m_K,K}^{(n_1)}$ for I_0. This is verified using (11) for the calculation of $|P_{m,K}^{(n_1)}|$. Security is then proved by applying Lemma 6 to find that for large enough k (and after enlarging K again if necessary), the distance between $\hat{x}(Kn_1 + (k-K)n_2)$ and $\hat{x}'(Kn_1 + (k-K)n_2)$ is lower-bounded by

$$\lambda^{Kn_1 + (k-K)n_2}\frac{\Omega}{\lambda-1}\left(\frac{L-1}{M_1-1} - \left(1 + 2\frac{\lambda-1}{\Omega}\right)\varepsilon\right).$$

C. Proof Sketch of Theorem 2

The elements of $\mathbf{A}^n$ that cannot be used as codewords are eliminated in successive steps. First, one eliminates every element $(a_t)_{t=1}^n$ that is the single element of a hyperedge in $H(\mathbf{T}_C^n)$ and considers the restriction $\mathbf{T}_C^n(1)$ of $\mathbf{T}_C^n$ to the reduced alphabet. Thus one obtains a new hypergraph $H(\mathbf{T}_C^n(1))$ and again eliminates all elements of the reduced alphabet which are the single element of a hyperedge of $H(\mathbf{T}_C^n(1))$. Continuing in this way, one eventually, say after $S^{(n)}$ steps, arrives at a reduced alphabet $\mathbf{A}_2^{(n)}$, having eliminated a set $\mathbf{A}_1^{(n)}$ of vertices. If we denote $\mathbf{T}_C^n$ restricted to $\mathbf{A}_2^{(n)}$ by $\mathbf{T}_C^n(S^{(n)})$, then all hyperedges of $H(\mathbf{T}_C^n(S^{(n)}))$ have at least two elements. Hence if we use every element of $\mathbf{A}_2^{(n)}$ as the codeword of a different message, the code $\mathbf{F}$ thus obtained is a zero-error wiretap $(|\mathbf{A}_2^{(n)}|, n)$-code.

Next one shows that $\mathbf{A}_1^{(n)} \subset (\mathbf{A}_1^{(1)})^n$, exploiting that $H(\mathbf{T}_C^n)$ is the n-fold square product of $H(\mathbf{T}_C)$ (the details can be found in [8]). Therefore $|\mathbf{A}_2^{(n)}| = |\mathbf{A}|^n - |\mathbf{A}_1^{(n)}| \geq |\mathbf{A}|^n - |\mathbf{A}_1^{(1)}|^n$, which grows in n at asymptotic exponential rate $\log|\mathbf{A}|$ if and only if $\mathbf{A}_2^{(1)} \subsetneq \mathbf{A}$. This proves Theorem 2.

REFERENCES

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

[2] 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.

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

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

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

[6] M. Hellmuth, L. Ostermeier, and P. Stadler, "A survey on hypergraph products," *Math. Comput. Sci.*, vol. 6, no. 1, pp. 1–32, 2012.

[7] M. van Dijk, "On a special class of broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. 43, no. 2, pp. 712–714, 1997.

[8] M. Wiese et al., "Uncertain Wiretap Channels and Secure Estimation", 2016. Available online at http://arxiv.org/abs/1605.00274