On the Strong Secrecy Capacity of Wiretap Channels with Side Information

Holger Boche and Rafael F. Schaefer

Lehrstuhl für Theoretische Informationstechnik
Technische Universität München, Germany

***Abstract*—The wiretap channel models the communication scenario where two legitimate users want to communicate in such a way that an external wiretapper is kept ignorant. In this paper the *wiretap channel with side information* is studied, where the wiretapper has additional side information about the transmitted message available for decoding. The corresponding secrecy capacity is derived and shown to be equal to the one of the classical wiretap channel without side information available at the wiretapper. Further, the capacity-achieving code structure is analyzed and properties are identified. Finally, channel uncertainty is taken into account and the results are extended to the compound wiretap channel with side information.**

I. Introduction

The concept of physical layer, or information theoretic, security is becoming more and more attractive, since it solely uses the physical properties of a wireless channel in order to establish security. So, regardless of what non-legitimate receivers do with the received signal, the confidential information cannot be reproduced with arbitrary high probability. Recently, there is growing interest in physical layer security; for instance see [1, 2] and references therein.

Physical layer security was initiated by Wyner, who introduced the *wiretap channel* [3]. It describes the simplest scenario involving security with one legitimate transmitter-receiver pair and one external wiretapper to be kept secret. The aim of the transmitter is to encode and transmit the confidential message in such a way that the legitimate receiver is able to decode the message and, at the same time, the wiretapper is prevented from inferring the confidential information from its received output. The wiretap channel is widely studied under several aspects, cf. for example [4–10].

Besides the wiretap channel, there is also work on multi-user settings such as the multiple access channel with confidential messages [11] or the MIMO Gaussian broadcast channel with common and confidential messages [12, 13].

All these works have in common that the wiretapper has only its received channel output available to infer on the confidential information. In this paper we study the *wiretap channel with side information*, where we consider a more powerful wiretapper which has additional side information about the transmitted message available. This additional side information models some a priori knowledge about the transmitted message which allows the wiretapper to restrict the message to a certain subset of all possible messages. Such side information can originate from previous transmissions due to certain network structures or from other cooperating wiretappers which share some knowledge with each other.

In Section II we introduce the wiretap channel with side information and discuss the model of side information at the wiretapper. In Section III we analyze the wiretap channel with side information in detail and derive the corresponding secrecy capacity region. It is shown that it is equal to one of the wiretap channel without side information, which means that additional side information cannot be exploited by the wiretapper. Further, the optimal code structure is characterized. In Section IV we take channel uncertainty into account applying the model of *compound channels* [14, 15]. We extend the results to the compound wiretap channel with side information and show that the results also hold in this case. But interestingly, one can show by proper examples based on [16] that this does not hold anymore if one considers arbitrarily varying channels. Finally, in Section V we conclude the paper.[1]

II. Wiretap Channel with Side Information

We start with some preliminaries and introduce the system model. Let $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ be finite input and output sets. Then for input and output sequences $x^n \in \mathcal{X}^n$, $y^n \in \mathcal{Y}^n$, and $z^n \in \mathcal{Z}^n$ of block length n, the discrete memoryless channels to the legitimate receiver and the wiretapper are given by $W^n(y^n|x^n) := \prod_{i=1}^n W(y_i|x_i)$ and $V^n(z^n|x^n) := \prod_{i=1}^n V(z_i|x_i)$, respectively. The wiretap channel is given by the pair of channels $\{W, V\}$ with common input.

A. Classical Wiretap Channel

In the classical wiretap channel, the task is to establish reliable communication between the transmitter and the legitimate receiver and, at the same time, keeping the confidential information secret from the wiretapper. Here, the wiretapper

This work was partly supported by the German Ministry of Education and Research (BMBF) under Grant 01BQ1050.

[1] *Notation:* Discrete random variables are denoted by capital letters and their realizations and ranges by lower case and script letters, respectively; $H(\cdot)$ and $I(\cdot;\cdot)$ are the traditional entropy and mutual information; $D(\cdot\|\cdot)$ is the Kullback-Leibler (information) divergence and $\|\mu - \nu\|$ is the total variation distance of measures μ and ν; $X - Y - Z$ denotes a Markov chain of random variables X, Y, and Z in this order; $\mathcal{P}(\cdot)$ is the set of all probability distributions.

978-1-4673-3122-7/13/$31.00 ©2013 IEEE

has only its channel output available to infer on the confidential communication. This is formalized as follows.

Definition 1: A (n, J_n)*-code* $\mathcal{C}_n$ for the wiretap channel consists of a stochastic encoder

$$E : \mathcal{J}_n \to \mathcal{P}(\mathcal{X}^n),$$

i.e., a stochastic matrix, with a set of messages $\mathcal{J}_n := \{1, ..., J_n\}$ and a collection of disjoint decoding sets

$$\big\{\mathcal{D}_j \subset \mathcal{Y}^n : j \in \mathcal{J}_n\big\}.$$

The average probability of error of a (n, J_n)-code $\mathcal{C}_n$ is given by

$$\bar{e}(\mathcal{J}_n) := \frac{1}{|\mathcal{J}_n|} \sum_{j \in \mathcal{J}_n} \sum_{x^n \in \mathcal{X}^n} W^n(\mathcal{D}_j^c | x^n) E(x^n | j).$$

Accordingly, the maximum probability of error is given by

$$e_{\max}(\mathcal{J}_n) := \max_{j \in \mathcal{J}_n} \sum_{x^n \in \mathcal{X}^n} W^n(\mathcal{D}_j^c | x^n) E(x^n | j). \tag{1}$$

To ensure that the message is kept secret from the wiretapper, we require

$$I(J; Z^n) \leq \epsilon_n \tag{2}$$

with J the random variable uniformly distributed over the set of messages $\mathcal{J}_n$ and $Z^n = (Z_1, Z_2, ..., Z_n)$ the output at the wiretapper. This criterion is known as *strong secrecy* [5].

Definition 2: A non-negative number R_S is an *achievable secrecy rate for the wiretap channel* if for all $\delta > 0$ there is an $n(\delta) \in \mathbb{N}$ and a sequence of (n, J_n)-codes $\{\mathcal{C}_n\}_{n \in \mathbb{N}}$ such that for all $n \geq n(\delta)$ we have $\frac{1}{n} \log J_n \geq R_S - \delta$ and

$$I(J; Z^n) \leq \epsilon_n$$

while $\bar{e}(\mathcal{J}_n) \to 0$ (or $e_{\max}(\mathcal{J}_n) \to 0$ respectively) and $\epsilon_n \to 0$ as $n \to \infty$. The *secrecy capacity* C_S is given by the supremum of all achievable secrecy rates R_S.

The secrecy capacity of the wiretap channel is well known, cf. for example [3–5, 10], and restated in the following theorem.

Theorem 1: The secrecy capacity C_S of the wiretap channel is given by

$$C_S = \max_{U - X - (Y,Z)} \big(I(U;Y) - I(U;Z)\big).$$

Remark 1: The seminal works [3, 4] considered the weak secrecy criterion for the wiretap channel. More recently, the secrecy capacity of the wiretap channel was also established for the strong secrecy criterion, cf. for example [5, 10].

B. Wiretap Channel with Side Information

In this paper we study more powerful wiretappers. More precisely, besides its channel output, the wiretapper has additional side information about the transmitted message available as depicted in Figure 1. Such side information can be based on a priori knowledge at the wiretapper, but can also originate from prior transmissions due to a certain network structure or from other cooperating wiretappers which help each other inferring the confidential communication.

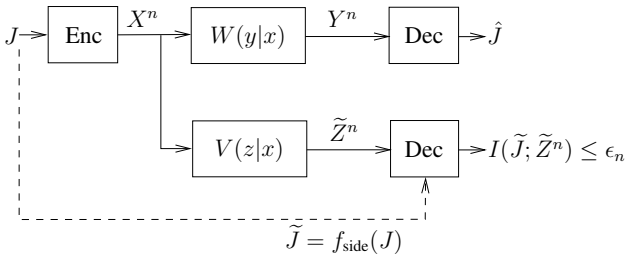

Fig. 1. Wiretap channel with side information at the wiretapper. The side information $f_{\text{side}}(J)$ available at the wiretapper restricts the transmitted message to the subset $\widetilde{\mathcal{J}} \subseteq \mathcal{J}_n$ with $|\widetilde{\mathcal{J}}| \geq 2$.

To model the side information available at the wiretapper, we introduce a deterministic function

$$f_{\text{side}} : \mathcal{J}_n \to \mathfrak{P}_2(\mathcal{J}_n)$$

with $\mathfrak{P}_2(\mathcal{J}_n)$ is the power set of all subsets of $\mathcal{J}_n$ with cardinality at least 2. So, for transmitted message $J \in \mathcal{J}_n$, the wiretapper is aware of $f_{\text{side}}(J) \in \mathfrak{P}_2(\mathcal{J}_n)$. Thus, the wiretapper can restrict the transmitted message to a subset $\widetilde{\mathcal{J}} \subseteq \mathcal{J}_n$ with $|\widetilde{\mathcal{J}}| \geq 2$. The restriction $|\widetilde{\mathcal{J}}| \geq 2$ is reasonable, since for $|\widetilde{\mathcal{J}}| = 1$ the transmitted message would be completely known at the wiretapper.

Remark 2: The case $|\widetilde{\mathcal{J}}| = 2$ models the scenario smallest uncertainty, since the wiretapper can restrict the transmitted message to only two alternatives.

Remark 3: Note that the scenario of no side information available at the wiretapper is included in the model by the special case $\widetilde{\mathcal{J}} = \mathcal{J}_n$.

Due to the side information at the wiretapper, the security requirement (2) slightly changes as follows. Now, we require that there is a universal ϵ_n (in the sense that it does not depend on the actual side information $\widetilde{\mathcal{J}}$) such that for all subsets $\widetilde{\mathcal{J}} \subseteq \mathcal{J}_n$ with $|\widetilde{\mathcal{J}}| \geq 2$ it holds

$$I(\widetilde{J}; \widetilde{Z}^n) \leq \epsilon_n \tag{3}$$

where $\widetilde{J}$ is the random variable uniformly distributed on the restricted set of messages $\widetilde{\mathcal{J}} \subseteq \mathcal{J}_n$ and $\widetilde{Z}^n = (\widetilde{Z}_1, \widetilde{Z}_2, ..., \widetilde{Z}_n)$ the corresponding output at the wiretapper.

Remark 4: It can be shown that the secrecy requirement (3) implies for the average decoding error $\bar{e}(\widetilde{\mathcal{J}})$ at the wiretapper

$$\bar{e}(\widetilde{\mathcal{J}}) \geq 1 - \frac{1}{|\widetilde{\mathcal{J}}|} - \lambda_n(\epsilon_n).$$

Thus, an important signal processing interpretation can be given to the information theoretic criterion (3). Regardless of the post-processing of the wiretapper, it always results in the worst behavior of decoding performance. For a detailed discussion we refer to [17].

Definition 3: A non-negative number R_S is an *achievable secrecy rate for the wiretap channel with side information* if for all $\delta > 0$ there is an $n(\delta) \in \mathbb{N}$, a universal ϵ_n, and a sequence of (n, J_n)-codes $\{\mathcal{C}_n\}_{n \in \mathbb{N}}$ such that for all $n \geq n(\delta)$ we have $\frac{1}{n} \log J_n \geq R_S - \delta$ and

$$I(\widetilde{J}; \widetilde{Z}^n) \leq \epsilon_n$$

for all subsets $\widetilde{\mathcal{J}} \subseteq \mathcal{J}_n$ with $|\widetilde{\mathcal{J}}| \geq 2$, while $\bar{e}(\mathcal{J}_n) \to 0$ (or $e_{\max}(\mathcal{J}_n) \to 0$ respectively) and $\epsilon_n \to 0$ as $n \to \infty$. The *secrecy capacity* $C_{S,\text{side}}$ is given by the supremum of all achievable secrecy rates R_S for the wiretap channel with side information.

For the analysis of the wiretap channel with side information, the following property of a wiretap code will be essential.

Definition 4: A code for the wiretap channel (with side information) has *vanishing output variation* if for all sets $\widetilde{\mathcal{J}} \subseteq \mathcal{J}_n$, there is a measure ϑ on $\mathcal{Z}^n$ such that for all $j \in \widetilde{\mathcal{J}}$ and

$$\overline{V}^n(z^n|j) \coloneqq \sum_{x^n \in \mathcal{X}^n} V^n(z^n|x^n) E(x^n|j)$$

it holds

$$\left\| \overline{V}^n(\cdot|j) - \vartheta \right\| \leq \epsilon_n$$

with $\epsilon_n \to 0$ as $n \to \infty$.

III. Secrecy Capacity Under Side Information

In this section we analyze the wiretap channel with side information in detail.

A. Secrecy Capacity

First, we analyze the secrecy capacity of the wiretap channel with side information and show that available side information at the wiretapper does not influence the secrecy capacity.

Theorem 2: The secrecy capacity of the wiretap channel with side information equals the secrecy capacity of the wiretap channel without side information, i.e.,

$$C_{S,\text{side}} = C_S.$$

Proof: To prove the desired result, we have to show the converse and achievability.

1) Proof of Converse: The inequality $C_{S,\text{side}} \leq C_S$ follows immediately, since additional side information at the wiretapper can only decrease the secrecy capacity. Thus, it remains to prove that C_S is actually achievable also in the case of side information available at the wiretapper.

2) Proof of Achievability: For the achievability we have to construct a wiretap code that realizes simultaneously reliable communication at the desired rate to the legitimate receiver, i.e., $e_{\max}(\mathcal{J}_n) \to 0$ as $n \to \infty$, and secrecy at the wiretapper, i.e., $I(\widetilde{J}; \widetilde{Z}^n) \to 0$ for all $\widetilde{\mathcal{J}} \subseteq \mathcal{J}_n$ with $|\widetilde{\mathcal{J}}| \geq 2$ as $n \to \infty$.

We know from [10] that there are wiretap codes with exponentially fast decreasing vanishing output variation, i.e., with ϵ_n in Definition 4 of the form $\epsilon_n = 2^{-n\beta}$ with $\beta > 0$, which achieve the secrecy capacity C_S of the wiretap channel (without side information). In particular, the maximum probability of error $e_{\max}(\mathcal{J}_n) \to 0$ as $n \to \infty$ for all $R_S < C_S$, which immediately implies for all subsets $\widetilde{\mathcal{J}} \subseteq \mathcal{J}_n$ with $|\widetilde{\mathcal{J}}| \geq 2$ that $e_{\max}(\widetilde{\mathcal{J}}) \to 0$ as well. Note that [10] treats the compound wiretap channel, but clearly, this code construction also works for the non-compound case.

Thus, we only have to check that this code also satisfies the security constraint at the wiretapper with side information. Therefore, let $\widetilde{\mathcal{J}}$ be an arbitrary set indicating the side information available at the wiretapper and $\widetilde{J}$ be the corresponding random variable uniformly distributed over $\widetilde{\mathcal{J}}$. Then the corresponding output distribution at the wiretapper is given by

$$\widetilde{V}^n(z^n) = \frac{1}{|\widetilde{\mathcal{J}}|} \sum_{j \in \widetilde{\mathcal{J}}} \overline{V}^n(z^n|j).$$

We denote the corresponding output random variable by $\widetilde{Z}^n$. From the vanishing output variation property of the wiretap code, cf. Definition 4, we know that there is a measure ϑ on $\mathcal{Z}^n$ such that for all $j \in \widetilde{\mathcal{J}}$ we have

$$\begin{aligned}
\left\| \overline{V}^n(\cdot|j) - \widetilde{V}^n(\cdot) \right\| &= \left\| \overline{V}^n(\cdot|j) - \vartheta + \vartheta - \widetilde{V}^n(\cdot) \right\| \\
&\leq \left\| \overline{V}^n(\cdot|j) - \vartheta \right\| + \left\| \vartheta - \widetilde{V}^n(\cdot) \right\| \\
&\leq \epsilon_n + \left\| \frac{1}{|\widetilde{\mathcal{J}}|} \sum_{j \in \widetilde{\mathcal{J}}} \left(\vartheta - \overline{V}^n(\cdot|j) \right) \right\| \\
&\leq \epsilon_n + \frac{1}{|\widetilde{\mathcal{J}}|} \sum_{j \in \widetilde{\mathcal{J}}} \left\| \vartheta - \overline{V}^n(\cdot|j) \right\| \\
&\leq 2\epsilon_n
\end{aligned}$$

with $\epsilon_n = 2^{-n\beta}$, $\beta > 0$, where the steps follow from the triangle inequality and the vanishing output variation property. Now, from the definition of mutual information and [18, Lemma 1.2.7], cf. also [10, Proof of Theorem 3.5], we get

$$\begin{aligned}
I(\widetilde{J}; \widetilde{Z}^n) &= \sum_{j \in \widetilde{\mathcal{J}}} \frac{1}{|\widetilde{\mathcal{J}}|} \left(H(\widetilde{V}^n(\cdot)) - H(\overline{V}^n(\cdot|j)) \right) \\
&= H(\widetilde{Z}^n) - H(\widetilde{Z}^n|\widetilde{J}) \\
&\leq -2\epsilon_n \log \frac{2\epsilon_n}{|\mathcal{Z}^n|} \\
&= -2\epsilon_n \log(2\epsilon_n) + 2n\epsilon_n \log|\mathcal{Z}| =: \epsilon'_n
\end{aligned}$$

with ϵ'_n independent of the actual set $\widetilde{\mathcal{J}} \subseteq \mathcal{J}_n$ and $\epsilon'_n \to 0$ as $n \to \infty$ since $\epsilon_n = 2^{-n\beta}$ decreases exponentially fast. ∎

This shows that side information at the wiretapper does not influence the secrecy capacity of the wiretap channel. Moreover, a code developed for the wiretap channel without side information is also suitable to protect the transmitted message in the event of additional side information at the wiretapper.

B. Optimal Code Structure

The previous analysis showed that a wiretap code with vanishing output variation is suitable to achieve the secrecy capacity of the wiretap channel with side information. Here we want to show that the reverse statement is also true. In more detail, the following result shows that an optimal code for the wiretap channel with side information has to have the vanishing output variation property.

Theorem 3: Let $\{\mathcal{C}_n\}_{n \in \mathbb{N}}$ be a sequence of wiretap codes achieving the secrecy capacity of the wiretap channel with side information. Let $E : \mathcal{J}_n \to \mathcal{P}(\mathcal{X}^n)$ be the corresponding stochastic encoder and $\overline{V}^n(z^n|j) \coloneqq \sum_{x^n \in \mathcal{X}^n} V^n(z^n|x^n) E(x^n|j)$. Then, there exists an $\epsilon_n =$

$2^{-n\beta}$ with some $\beta > 0$ and a measure ϑ on $\mathcal{Z}^n$ such that for all $j \in \mathcal{J}_n$ it holds

$$\left\|\overline{V}^n(\cdot|j) - \vartheta\right\| \leq \epsilon_n.$$

This means, the optimal code has the vanishing output variation property.

Proof: Let $\widetilde{\mathcal{J}} = \{j_1, j_2\} \subseteq \mathcal{J}_n$ be an arbitrary message subset with two elements. Since the code is optimal for the wiretap channel with side information by assumption, we have

$$I(\widetilde{J}; \widetilde{Z}^n) \leq \epsilon_n \tag{4}$$

with ϵ_n independent of $\widetilde{\mathcal{J}}$. Let $P_{\widetilde{J}\widetilde{Z}^n}(j, z^n) = \overline{V}^n(z^n|j)P_{\widetilde{J}}(j)$ be the joint distribution and $P_{\widetilde{J}}$ and $P_{\widetilde{Z}^n}$ be the corresponding marginals.

Writing (4) in terms of Kullback-Leibler (information) divergence as

$$I(\widetilde{J}; \widetilde{Z}^n) = D(P_{\widetilde{J}\widetilde{Z}^n} \| P_{\widetilde{J}} \otimes P_{\widetilde{Z}^n}) \leq \epsilon_n$$

with $P_{\widetilde{J}} \otimes P_{\widetilde{Z}^n}(j, z^n) = P_{\widetilde{J}}(j)P_{\widetilde{Z}^n}(z^n)$ for all $j \in \widetilde{\mathcal{J}}$ and $z^n \in \mathcal{Z}^n$, we get by Pinsker's inequality, cf. for example [18, Problem 3.18], for some constant $c > 0$ the following

$$\begin{aligned}
c\sqrt{\epsilon_n} &\geq \left\|P_{\widetilde{J}\widetilde{Z}^n} - P_{\widetilde{J}} \otimes P_{\widetilde{Z}^n}\right\| \\
&= \sum_{z^n \in \mathcal{Z}^n} \left|P_{\widetilde{J}\widetilde{Z}^n}(j_1, z^n) - P_{\widetilde{J}}(j_1)P_{\widetilde{Z}^n}(z^n)\right| \\
&\qquad + \sum_{z^n \in \mathcal{Z}^n} \left|P_{\widetilde{J}\widetilde{Z}^n}(j_2, z^n) - P_{\widetilde{J}}(j_2)P_{\widetilde{Z}^n}(z^n)\right| \\
&= \sum_{z^n \in \mathcal{Z}^n} \left|\frac{1}{2}\overline{V}^n(z^n|j_1) - \frac{1}{2}P_{\widetilde{Z}^n}(z^n)\right| \\
&\qquad + \sum_{z^n \in \mathcal{Z}^n} \left|\frac{1}{2}\overline{V}^n(z^n|j_2) - \frac{1}{2}P_{\widetilde{Z}^n}(z^n)\right|.
\end{aligned}$$

Here, the first equality follows from the definition of total variation distance and the fact that $\widetilde{\mathcal{J}}$ has only two elements, and the second equality from the fact that $\widetilde{J}$ is uniformly distributed. Thus, for each $l = 1, 2$ we have

$$\sum_{z^n \in \mathcal{Z}^n} \left|\overline{V}^n(z^n|j_l) - P_{\widetilde{Z}^n}(z^n)\right| \leq 2c\sqrt{\epsilon_n}.$$

Since $P_{\widetilde{Z}^n}(z^n) = \frac{1}{2}\big(\overline{V}^n(z^n|j_1) + \overline{V}^n(z^n|j_2)\big)$, we have

$$\begin{aligned}
&\sum_{z^n \in \mathcal{Z}^n} \left|\overline{V}^n(z^n|j_1) - \frac{1}{2}\overline{V}^n(z^n|j_1) - \frac{1}{2}\overline{V}^n(z^n|j_2)\right| \\
&\qquad = \frac{1}{2}\sum_{z^n \in \mathcal{Z}^n} \left|\overline{V}^n(z^n|j_1) - \overline{V}^n(z^n|j_2)\right| \leq 2c\sqrt{\epsilon_n}.
\end{aligned}$$

Thus, by the definition of total variation distance we have for arbitrary $j_1, j_2 \in \mathcal{J}_n$

$$\begin{aligned}
&\sum_{z^n \in \mathcal{Z}^n} \left|\overline{V}^n(z^n|j_1) - \overline{V}^n(z^n|j_2)\right| \\
&\qquad = \left\|\overline{V}^n(\cdot|j_1) - \overline{V}^n(\cdot|j_2)\right\| \leq 4c\sqrt{\epsilon_n}.
\end{aligned}$$

Now, we set

$$\vartheta(z^n) = \frac{1}{J_n}\sum_{j \in \mathcal{J}_n} \overline{V}^n(z^n|j)$$

for all $z^n \in \mathcal{Z}^n$, so that for any $l \in \mathcal{J}_n$ we have

$$\begin{aligned}
\left\|\overline{V}^n(\cdot|l) - \vartheta\right\| &= \left\|\overline{V}^n(\cdot|l) - \frac{1}{J_n}\sum_{j \in \mathcal{J}_n} \overline{V}^n(\cdot|j)\right\| \\
&= \left\|\frac{1}{J_n}\sum_{j \in \mathcal{J}_n} \big(\overline{V}^n(\cdot|l) - \overline{V}^n(\cdot|j)\big)\right\| \\
&\leq \frac{1}{J_n}\sum_{j \in \mathcal{J}_n} \left\|\overline{V}^n(\cdot|l) - \overline{V}^n(\cdot|j)\right\| \leq 4c\sqrt{\epsilon_n}.
\end{aligned}$$

This means an optimal code for the wiretap channel with side information always has the vanishing output variation property. This completes the proof of the theorem. ∎

IV. Compound Wiretap Channel

In the previous analysis we assumed that transmitter and receiver have perfect channel state information (CSI). But in practical systems there is always uncertainty in channel state information due to the nature of the wireless medium but also due to implementational issues such as imperfect channel estimation or limited feedback schemes. A reasonable model is to assume that the exact channel realization is not known; rather, it is only known that it belongs to a pre-specified set of channels. If this channel remains fixed during the whole transmission of a codeword, this corresponds to the concept of *compound channels* [14, 15].

To model the compound wiretap channel, let $\mathcal{T}$ be a finite index set. Then for fixed $t \in \mathcal{T}$, the discrete memoryless channels to the legitimate receiver and the wiretapper are given by $W_t^n(y^n|x^n) \coloneqq \prod_{i=1}^n W_t(y_i|x_i)$ and $V_t^n(z^n|x^n) \coloneqq \prod_{i=1}^n V_t(z_i|x_i)$.

Definition 5: The discrete memoryless *compound wiretap channel* $\mathfrak{W}$ is given by $\mathfrak{W} \coloneqq \{(W_t, V_t) : t \in \mathcal{T}\}$.

Since transmitter and receiver have only imperfect CSI and do not know the exact channel realization, they have to use universal encoder and decoder which do not depend on the actual channel realization similarly as in Definition 1. The only difference is that we have to ensure that the legitimate receiver can decode the transmitted message for all channel realizations $t \in \mathcal{T}$ so that the average probability and maximum probability of error, cf. (1), become $\bar{e}(\mathcal{J}_n) \coloneqq \max_{t \in \mathcal{T}} \frac{1}{|\mathcal{J}_n|}\sum_{j \in \mathcal{J}_n}\sum_{x^n \in \mathcal{X}^n} W_t^n(\mathcal{D}_j^c|x^n)E(x^n|j)$. and

$$e_{\max}(\mathcal{J}_n) \coloneqq \max_{t \in \mathcal{T}} \max_{j \in \mathcal{J}_n} \sum_{x^n \in \mathcal{X}^n} W_t^n(\mathcal{D}_j^c|x^n)E(x^n|j).$$

Accordingly, since the exact channel realization to the wiretapper is not known, we have to ensure that the message is kept secret for all possible channel realizations $t \in \mathcal{T}$. Thus, the security requirement in (3) reads as

$$\max_{t \in \mathcal{T}} I(\widetilde{J}; \widetilde{Z}_t^n) \leq \epsilon_n$$

for all subsets $\widetilde{\mathcal{J}} \subseteq \mathcal{J}_n$ with $|\widetilde{\mathcal{J}}| \geq 2$ and ϵ_n independent of $\widetilde{\mathcal{J}}$. With these expressions the definition of an *achievable secrecy rate for the compound wiretap channel with side information* and the corresponding *secrecy capacity* $C_{S,\text{side}}(\mathfrak{W})$ follow accordingly.

To establish similar results as for the case with perfect CSI, we further need the vanishing output variation property which is slightly adapted to the compound setting as follows.

Definition 6: A code for the compound wiretap channel (with side information) has *vanishing output variation* if for all sets $\widetilde{\mathcal{J}} \subseteq \mathcal{J}_n$ there are measures ϑ_t, $t \in \mathcal{T}$, such that for all $j \in \widetilde{\mathcal{J}}$ and

$$\overline{V}_t^n(z^n|j) \coloneqq \sum_{x^n \in \mathcal{X}^n} V_t^n(z^n|x^n) E(x^n|j)$$

it holds

$$\left\| \overline{V}_t^n(\cdot|j) - \vartheta_t \right\| \leq \epsilon_n$$

with $\epsilon_n \to 0$ as $n \to \infty$.

The difference to Definition 4 is that it is sufficient to have for each channel realization $t \in \mathcal{T}$ a different measure ϑ_t. This is motivated by [10] and we obtain the following result.

Theorem 4: An achievable secrecy rate R_S for the compound wiretap channel with side information is given by

$$R_S = \max_{P_X \in \mathcal{P}(\mathcal{X})} \left(\min_{t \in \mathcal{T}} I(X;Y_t) - \max_{t \in \mathcal{T}} I(X;Z_t) \right) \qquad (5)$$

with Y_t and Z_t the corresponding output random variables for channel realization $t \in \mathcal{T}$.

Sketch of Proof: Similarly as in the proof of Theorem 1 we can make use of the code construction given in [10]. In more detail, it already presents the construction of a code for the wiretap channel (without side information) which has the vanishing output variation property as stated in Definition 6. As in Theorem 2 it is straight forward to show that this code achieves the rate stated in (5) establishing the reliability part of the proof.

Similarly as in the proof of Theorem 2 it remains to check, if the security requirement is fulfilled. The code has the vanishing output variation property suitable for the compound wiretap channel, cf. Definition 6, and in particular, the measure ϑ_t can be different for each realization $t \in \mathcal{T}$. With this, the proof is analogous to the proof given in Theorem 1 and omitted for brevity. ∎

Accordingly, it is straight forward to show, similarly as in Theorem 3, that for the compound wiretap channel with side information, the optimal code must have the vanishing output variation property.

Corollary 1: Let $\{\mathcal{C}_n\}_{n \in \mathbb{N}}$ be a sequence of wiretap codes achieving the secrecy capacity of the compound wiretap channel with side information. Let $E : \mathcal{J}_n \to \mathcal{P}(\mathcal{X}^n)$ be the corresponding stochastic encoder and $\overline{V}_t^n(z^n|j) \coloneqq \sum_{x^n \in \mathcal{X}^n} V_t^n(z^n|x^n) E(x^n|j)$, $t \in \mathcal{T}$. Then, there exists an $\epsilon_n = 2^{-n\beta}$ with some $\beta > 0$ and measures ϑ_t, $t \in \mathcal{T}$, such that for all $j \in \mathcal{J}_n$ it holds

$$\left\| \overline{V}_t^n(\cdot|j) - \vartheta_t \right\| \leq \epsilon_n$$

for all $t \in \mathcal{T}$. This means, the optimal code has the almost vanishing output variation property, cf. Definition 6.

V. Concluding Remarks

Previous works focused on wiretappers which solely use their received channel output to infer on the confidential information. In this paper we analyzed more powerful wiretappers with additional side information. Interestingly, side information at the wiretapper does not affect the secrecy capacity. It is shown that the secrecy capacity of the wiretap channel with side information equals the one without side information. Moreover, this carries over to compound channels, where again, the side information does not affect the achievable rates for the compound wiretap channel with side information. But interestingly, it can easily be shown by proper examples based on [16] that this does not hold if one consider arbitrarily varying wiretap channels. Here, the secrecy capacities of the case with and without side information are different in general.

References

[1] Y. Liang, H. V. Poor, and S. Shamai (Shitz), "Information Theoretic Security," *Foundations and Trends in Communications and Information Theory*, vol. 5, no. 4-5, pp. 355–580, 2009.

[2] M. Bloch and J. Barros, *Physical-Layer Security: From Information Theory to Security Engineering*. Cambridge University Press, 2011.

[3] A. D. Wyner, "The Wire-Tap Channel," *Bell Syst. Tech. J.*, vol. 54, pp. 1355–1387, Oct. 1975.

[4] I. Csiszár and J. Körner, "Broadcast Channels with Confidential Messages," *IEEE Trans. Inf. Theory*, vol. 24, no. 3, pp. 339–348, May 1978.

[5] U. M. Maurer and S. Wolf, "Information-Theoretic Key Agreement: From Weak to Strong Secrecy for Free," in *EUROCRYPT 2000, Lecture Notes in Computer Science*. Springer-Verlag, May 2000, vol. 1807, pp. 351–368.

[6] Y. Liang, G. Kramer, H. V. Poor, and S. Shamai (Shitz), "Compound Wiretap Channels," *EURASIP J. Wireless Commun. Netw.*, vol. Article ID 142374, pp. 1–13, 2009.

[7] A. Khisti and G. W. Wornell, "Secure Transmission With Multiple Antennas I: The MISOME Wiretap Channel," *IEEE Trans. Inf. Theory*, vol. 56, no. 7, pp. 3088–3104, Jul. 2010.

[8] ——, "Secure Transmission With Multiple Antennas–Part II: The MIMOME Wiretap Channel," *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5515–5532, Nov. 2010.

[9] E. Ekrem and S. Ulukus, "On Gaussian MIMO Compound Wiretap Channels," in *Proc. Conf. Inf. Sciences and Systems*, Baltimore, MD, USA, Mar. 2010, pp. 1–6.

[10] I. Bjelaković, H. Boche, and J. Sommerfeld, "Secrecy Results for Compound Wiretap Channels," *Probl. Inf. Transmission*, 2013, accepted.

[11] Y. Liang and H. V. Poor, "Multiple-Access Channels With Confidential Messages," *IEEE Trans. Inf. Theory*, vol. 54, no. 3, pp. 976–1002, Mar. 2008.

[12] R. Liu, T. Liu, H. V. Poor, and S. Shamai (Shitz), "MIMO Gaussian Broadcast Channels with Confidential and Common Messages," in *Proc. IEEE Int. Symp. Inf. Theory*, Austin, TX, USA, Jun. 2010, pp. 2578–2582.

[13] E. Ekrem and S. Ulukus, "Gaussian MIMO Broadcast Channels with Common and Confidential Messages," in *Proc. IEEE Int. Symp. Inf. Theory*, Austin, TX, USA, Jun. 2010, pp. 2583–2587.

[14] D. Blackwell, L. Breiman, and A. J. Thomasian, "The Capacity of a Class of Channels," *Ann. Math. Stat.*, vol. 30, no. 4, pp. 1229–1241, Dec. 1959.

[15] J. Wolfowitz, "Simultaneous Channels," *Arch. Rational Mech. Analysis*, vol. 4, no. 4, pp. 371–386, 1960.

[16] I. Bjelaković, H. Boche, and J. Sommerfeld, "Strong Secrecy in Arbitrarily Varying Wiretap Channels," in *Proc. IEEE Inf. Theory Workshop*, Lausanne, Switzerland, Sep. 2012.

[17] H. Boche and R. F. Wyrembelski, "Wiretap Channels with Side Information Strong Secrecy Capacity and Optimal Transceiver Design," 2012, submitted for publication.

[18] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*, 2nd ed. Cambridge University Press, 2011.