INVITED PAPER

Secure Communication Under Channel Uncertainty and Adversarial Attacks

This paper surveys various channel models that capture channel uncertainty at transmitters and adversarial attacks, and reviews design of information-theoretic schemes to achieve secure communications and characterization of the secrecy capacity for these models.

By Rafael F. Schaefer, *Member IEEE*, Holger Boche, *Fellow IEEE*, and H. Vincent Poor, *Fellow IEEE*

ABSTRACT | Information theoretic approaches to security have been examined as a promising complement to current cryptographic techniques. Such information theoretic approaches establish reliable communication and data confidentiality directly at the physical layer of a communication network by taking the properties of the noisy channel into account leading to unconditional security regardless of the computational capabilities of eavesdroppers. The provision of accurate channel state information is a major challenge particularly in wireless communication systems, especially information about the channels to eavesdroppers. In addition, there might be malevolent adversaries who jam or influence the channel of the legitimate users. This paper surveys different models for secure communication under channel uncertainty and adversarial attacks and reviews the corresponding secrecy capacity results, which characterize the maximum rate at which information can be sent to legitimate receivers while being kept perfectly security from eavesdroppers.

KEYWORDS | Arbitrarily varying channel; common randomness; compound channel; continuity; robustness; secrecy capacity; wiretap channel

Manuscript received January 1, 2015; revised April 23, 2015; accepted July 15, 2015. Date of publication August 21, 2015; date of current version September 16, 2015. The work of R. F. Schaefer was supported by the German Research Foundation (DFG) under Grant WY 151/2-1. The work of H. Boche was supported in part by the German Research Foundation (DFG) under Grant BO 1734/20-1, and in part by the German Ministry of Education and Research (BMBF) under Grants 01BQ1050 and 16KIS0118. The work of H. V. Poor was supported by the U.S. National Science Foundation under Grant CMMI-1435778.

R. F. Schaefer and **H. V. Poor** are with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544 USA (e-mail: rafaelfs@princeton.edu; poor@princeton.edu).

H. Boche is with the Lehrstuhl für Theoretische Informationstechnik, Technische Universität München, 80290 München, Germany (e-mail: boche@tum.de).

Digital Object Identifier: 10.1109/JPROC.2015.2459652

I. INTRODUCTION

Rapid developments in communication systems make information available almost everywhere at any time. Along with this, the security of sensitive information from unauthorized access becomes an important issue for the design of such systems. This is in particular crucial for wireless communication systems as they are inherently vulnerable to eavesdropping: Due to the open nature of the wireless medium, transmitted signals are received not only by the intended users but also are easily eavesdropped upon by non-legitimate receivers.

The architecture of current communication systems usually separates error correction and data encryption. The former is typically realized at the physical layer, transforming the noisy communication channel into a reliable "bit pipe." Then the data encryption is implemented on top of that by applying cryptographic principles. A drawback of this approach is that it relies on the assumption of insufficient computational capabilities of non-legitimate receivers resulting in so-called conditional security.

In recent years, *information theoretic approaches to security* have been intensively examined as a complement to cryptographic techniques. Such approaches establish reliable communication and data confidentiality jointly at the physical layer by taking the properties of the communication channel into account. The first work in this area goes back to Shannon, who showed in his seminal paper [1] that a secret key used as a *one-time pad* allows for secure communication over a noiseless channel. Subsequently, Wyner introduced the *wiretap channel* in [2], which describes the communication scenario over a noisy channel and without secret keys. In this context, he introduced the notion of *secrecy capacity*, which is defined as the maximum rate at

which information can be sent to a legitimate receiver while being kept perfectly security from an eavesdropper. Later, this framework was generalized by Csiszár and Körner to the *broadcast channel with confidential messages (BCC)* [3]. Recently, this area has drawn considerable attention since it provides a promising approach to achieve unconditional security regardless of the computational capabilities of non-legitimate receivers; see for example [4]–[8] and references therein. Concurrently, it has been demonstrated that secure communication can efficiently be embedded into wireless networks by jointly implementing it with other non-secure services at the physical layer [9]. Thus, it is not surprising that operators of wireless communication systems and national agencies have also identified this concept as a key technique to secure future communication systems [10]–[12].

Many of the initial studies in the area of information theoretic security have in common that all channels (including those to non-legitimate eavesdroppers) are assumed to be perfectly known to all users and fixed during the entire duration of transmission. This is termed *perfect channel state information (CSI)* and such idealized communication conditions allow one to obtain an understanding and important insights of the fundamental principles of information theoretic security. These are briefly reviewed and discussed in Section II. In particular, the secrecy capacity of the wiretap channel has been established for discrete memoryless channels in [2], [3], [13], and [14] and for multiple-input multiple-output (MIMO) Gaussian channels in [15]–[18].

However, in practical systems CSI will always be limited due to the nature of the wireless medium and estimation/feedback inaccuracy. In addition, malevolent eavesdroppers will not provide any information about their channels to legitimate users which makes the assumption of perfect eavesdropper CSI questionable. Accordingly, limited CSI (especially to potential eavesdroppers) must be assumed to ensure reliability and data confidentiality in a robust way.

A first step in the direction of more realistic and practically relevant CSI assumptions is given by the concept of a *compound channel* [19], [20]. In this model, the actual channel realization is unknown. Rather, it is only known to the users that the true channel realization belongs to a known set of channels (uncertainty set) and that it remains constant for the whole duration of transmission. Accordingly, the *compound wiretap channel* models secure communication over compound channels and has been studied in [21]–[29]. Despite these efforts, a single-letter description of the secrecy capacity is known only for special cases such as degraded channels [21], [22] or certain multiple-input multiple-output (MIMO) Gaussian channels [25]–[28]. However, a single-letter characterization of the secrecy capacity that holds for the general case remains unknown to date (if it exists at all). Only a multi-letter description of the secrecy capacity has been established so far [22]. This is discussed in detail in Section III.

While for compound channels the unknown channel realization remains constant for the entire duration of transmission, the concept of an *arbitrarily varying channel (AVC)* [30]–[32] provides a model in which this realization may vary from channel use to channel use in an unknown and arbitrary manner. The corresponding *arbitrarily varying wiretap channel (AVWC)* has been studied in [33]–[40] and it has been shown that it makes a difference whether unassisted or common randomness (CR) assisted codes are used by the transmitter and legitimate receiver. In particular, if the channel to the legitimate receiver possesses the so-called property of symmetrizability, the unassisted secrecy capacity is zero, while the CR-assisted secrecy capacity may be non-zero. A complete characterization of the relation between the unassisted and CR-assisted secrecy capacity has been established in [34] and [38]; but similar to the compound wiretap channel, a single-letter characterization of the secrecy capacity itself remains open. CR-assisted achievable secrecy rates are known only under certain circumstances [33], [34], [36]. Recently, a multi-letter description of the CR-assisted secrecy capacity has been found in [37]. This is the content of Section IV.

Whenever a communication system is composed of orthogonal sub-systems, its overall performance is determined by the sum of all sub-systems. Taking an orthogonal frequency division multiplexing (OFDM) system as an example, its overall capacity is given by the sum of the capacities of all sub-channels. To this end, a system consisting of two orthogonal ordinary channels, where both are "useless" in the sense of having zero capacity, the overall capacity of the system is zero as well. This reflects the world view of classical additivity of resources in the sense that "$0 + 0 = 0$." In contrast to that, studies have revealed surprising phenomena for secure communication over AVCs: Two orthogonal AVWCs, each useless by itself in the sense that it has a zero unassisted secrecy capacity, can be used together to super-activate the whole system to allow for secure communication at non-zero secrecy rates [35], [38], [40]. This shows that the classical additivity of orthogonal resources does not hold anymore when secrecy requirements are imposed (in the sense that "$0 + 0 > 0$"). To date, such phenomena have been observed only for quantum communication systems [41], [42] and to the best of our knowledge, this is the first example that such effects can happen for classical communication systems as well.

So far the concepts of compound and arbitrarily varying wiretap channels are motivated from a channel uncertainty point of view. But these concepts are also suitable to model secure communication in the presence of active adversaries and their potential attacks on the communication. For example, a malicious adversary may influence or jam the legitimate channel by choosing the channel realization that governs the transmission. In such a case, the transmitter and legitimate receiver have to design their

encoding-decoding functions universally, since they usually have no knowledge about the strategy or the intentions of the adversary and therewith no knowledge about the expected channel realization. Thus, such attacks are perfectly modeled by compound wiretap channels. Accordingly, AVWCs would then model even more powerful adversaries, whose jamming strategies change with time. Again, the absence of any knowledge requires the legitimate users to prepare for a channel that may vary in an unknown and arbitrary manner from channel use to channel use. This has been done in [35], where the optimal jamming strategy of the adversary has been identified and it is shown that it differs depending on whether the adversary has access to the common randomness or not. A related problem is the so-called covert communication in which the legitimate users wish to communicate in such a way that the whole communication itself is not detectable by non-legitimate eavesdroppers. This is another potential strategy for the legitimate users to ensure secrecy and even avoid adversarial attacks [43]–[45].

Studying secure communication from an adversarial perspective reveals the following obvious observation: the secrecy capacities of the compound and arbitrarily varying wiretap channels depend on the underlying uncertainty set. In general, the performance of a communication system should depend in a continuous way on its system parameters. Since, if small changes in the parameters would lead to dramatic losses in performance, the corresponding approach is not robust and will most likely not be used. This means in the context of secure communication that the secrecy capacity should depend in a continuous way on the underlying uncertainty set. Then approaches are desirable that are robust against variations so that small variations in the uncertainty set result in small variations in the corresponding secrecy capacity. Such a continuous dependency is in particular desirable in the context of active adversaries who can influence the system parameters in a malicious way. In [39] it has been shown that the secrecy capacity of the compound wiretap channel possesses this behavior of a continuous dependence on the uncertainty set, while for AVWCs this might not be the case anymore. Here, it might happen that small changes in the uncertainty set lead to a dramatic loss in secrecy capacity. This line of study is continued in [39] and [46], where it has been shown that not only does the secrecy capacity possess this behavior but so do the corresponding codes themselves.

Up to this point, the simplest model of secure communication has been discussed: a single transmitter-receiver pair in the presence of one external eavesdropper. There has been some effort to extend the previously discussed concepts to more complex multi-user scenarios as well. Most noteworthy in this context is the broadcast channel with confidential messages [3]. Similar to the wiretap channel, the sender transmits a confidential message to a legitimate receiver while keeping an eavesdropper ignorant. Additionally, the sender transmits a common message as well which is intended for both the legitimate receiver and the eavesdropper. Thus, in this case, the eavesdropper has two different roles: It is a legitimate receiver of the common message and a non-legitimate receiver of the confidential message. This scenario has been studied in [47]–[49] for compound channels. Unfortunately, similar to the wiretap channel only a multi-letter characterization of the secrecy capacity region is known and only achievable secrecy rate regions have been established in a single-letter version. Besides the compound BCC, there are initial studies for the compound multiple access wiretap channel [50] and compound multiple access channel with confidential messages [51].

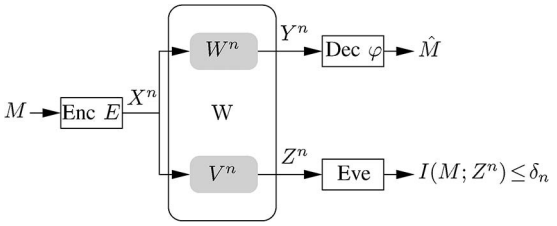

Fig. 1. *Wiretap channel W. The transmitter encodes the message M into the codeword $X^n = E(M)$ and transmits it over the wiretap channel W to the legitimate receiver, which has to decode its intended message $\hat{M} = \varphi(Y^n)$. At the same time, the eavesdropper has to be kept ignorant of M in the sense that $I(M;Z^n) \leq \delta_n$ must hold.*

II. WIRETAP CHANNEL UNDER PERFECT CSI

In this section we introduce the *wiretap channel* which is the simplest scenario involving security with one legitimate transmitter-receiver pair and one eavesdropper to be kept ignorant of the transmitted message as shown in Fig. 1. We start with the ideal assumption of perfect CSI at all users and present the basic ideas and concepts of secure communication over such a wiretap channel. The results and insights of the perfect CSI case will then allow us to approach the case of imperfect CSI later.

A. System Model

Throughout this paper, we assume that the inputs and outputs are from finite alphabets. This is motivated by the fact that a transmitter must use a finite modulation scheme such as BPSK or QAM due to practical limitations. Then on the receiver side, a received signal must be quantized before further digital processing. Thus, it is reasonable to assume finite input and output sets denoted by $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ in the following.

A perfectly known channel between the transmitter and the legitimate receiver can then be expressed by a stochastic matrix $W : \mathcal{X} \to \mathcal{P}(\mathcal{Y})$. Such a matrix describes

the probabilistic mapping between the input symbols and their potential outputs symbols received by the legitimate receiver. Assuming a *discrete memoryless channel (DMC)*, the probability law for a transmission of block length n is

$$W^n(y^n|x^n) = \prod_{i=1}^{n} W(y_i|x_i)$$

with input sequence $x^n = (x_1, x_2, \ldots, x_n) \in \mathcal{X}^n$ and output sequence $y^n = (y_1, y_2, \ldots, y_n) \in \mathcal{Y}^n$. Here x_i and y_i represent the channel input and output at time instant i, $i = 1, 2, \ldots, n$. Thus, for DMCs the current output depends only on the current input and not on previous inputs.

Similarly, we can model the channel between the transmitter and the eavesdropper by a stochastic matrix $V : \mathcal{X} \to \mathcal{P}(\mathcal{Z})$ and assume $V^n(z^n|x^n) = \prod_{i=1}^{n} V(z_i|x_i)$. Having specified the channels to the legitimate receiver and the eavesdropper, we define the wiretap channel as follows.

Definition 1: The discrete memoryless *wiretap channel* W is given by the pair of channels with common inputs as

$$\mathrm{W} = \{W, V\}.$$

Note that since the legitimate receiver and eavesdropper are not supposed to cooperate, there is no loss in generality by representing a wiretap channel by its marginal probabilities and not by its joint probability distribution. As a consequence, in this framework two wiretap channels with different joint probability distributions, but the same marginals will lead to the same secrecy capacity, cf. for example [4, Lemma 2.1]. This is because the performance of a code is measured based on the marginal distributions and not on the joint one as we will see in the next subsection.

B. Wiretap Codes

The communication task for the wiretap channel is twofold: Reliable communication between the transmitter and the legitimate receiver must be enabled and, at the same time, this communication must be kept secret from the eavesdropper. Important in this setup is that the eavesdropper has only its own channel output available to infer the confidential information. This is formalized as follows.

Definition 2: An (n, M_n)*-code* $\mathcal{C}$ for the wiretap channel consists of one stochastic encoder at the transmitter

$$E : \mathcal{M} \to \mathcal{P}(\mathcal{X}^n) \tag{1}$$

i.e., a stochastic matrix, with a set of confidential messages $\mathcal{M} = \{1, \ldots, M_n\}$ and a deterministic decoder at the legitimate receiver

$$\varphi : \mathcal{Y}^n \to \mathcal{M}. \tag{2}$$

The rate of this code is defined as $(1/n) \log M_n$.

The traditional approach is to use a deterministic encoder at the transmitter, which is a one-to-one mapping assigning each message $m \in \mathcal{M}$ exactly one codeword $x^n \in \mathcal{X}^n$. Such an encoder is sufficient for communication scenarios in which no security constraints are imposed. However, it has been shown that this is no longer true in the presence of eavesdroppers and the encoder E in (1) then needs to be stochastic. This means that it is specified by conditional probabilities $E(x^n|m)$ with $\sum_{x^n \in \mathcal{X}^n} E(x^n|m) = 1$ for each $m \in \mathcal{M}$, where $E(x^n|m)$ is the probability that the message $m \in \mathcal{M}$ is encoded as $x^n \in \mathcal{X}^n$. On the other hand, there is no benefit in using a stochastic decoder instead of deterministic one in (2), cf. for example [7, Sec. 3.4].

The quality of such a code is measured by two performance criteria: reliability and security. The reliability criterion ensures that the legitimate receiver is always able to decode its intended message. When the transmitter has sent the message $m \in \mathcal{M}$ and the legitimate receiver has received the channel output $y^n \in \mathcal{Y}^n$, it makes a decoding error if $\varphi(y^n) \neq m$. Thus, the probability of error for message $m \in \mathcal{M}$ is given by

$$e(m) = \sum_{x^n \in \mathcal{X}^n} \sum_{y^n : \varphi(y^n) \neq m} W^n(y^n|x^n) E(x^n|m).$$

Assuming all messages to be uniformly distributed lead to the average probability of error criterion

$$\bar{e} = \frac{1}{|\mathcal{M}|} \sum_{m \in \mathcal{M}} e(m). \tag{3}$$

A stronger notion of reliability is given by the maximum probability of error criterion

$$e_{\max} = \max_{m \in \mathcal{M}} e(m). \tag{4}$$

Obviously, the maximum probability of error (4) is a stronger performance criterion than the average probability of error (3). For perfect CSI both criteria will lead for the wiretap channel to the same secrecy capacity [3]. However, this is not always the case as for certain channels

these two criteria can result in different capacities. The second performance measure of security is discussed in detail in the following.

C. Secrecy Criterion

The secrecy criterion ensures that the non-legitimate eavesdropper is not able to infer any information about the transmitted message. Let M be a random variable uniformly distributed over the set of messages $\mathcal{M}$ and $Z^n = (Z_1, Z_2, \ldots, Z_n)$ be the channel output at the eavesdropper, see also Fig. 1. In his seminal paper [2], Wyner defined the secrecy of the confidential message in terms of equivocation having in mind that the channel output at the eavesdropper Z^n should not reveal any information about the message M. He required $(1/n)H(M|Z^n) \approx (1/n)H(M)$ so that the channel output Z^n does not decrease the uncertainty about the transmitted message M in terms of "rate" (due to the factor $1/n$). This criterion has been termed *weak secrecy* and is often equivalently written as

$$\frac{1}{n} I(M; Z^n) \leq \delta_n \tag{5}$$

with $\delta_n \to 0$ as $n \to \infty$. Recently, this criterion has been strengthened by dropping the division by the block length n as

$$I(M; Z^n) \leq \delta_n \tag{6}$$

with $\delta_n \to 0$ as $n \to \infty$. This is known as *strong secrecy* and the intuition is to have the total amount of information leaked to the eavesdropper small. Strong secrecy for the wiretap channel was first considered in [13] and [14]. Recently, different approaches have been proposed to achieve strong secrecy [52]–[54].

Vanishing information leakage implies that the average probability of error $\bar{e}_{\text{Eve}}$, as in (3), at the eavesdropper approaches one as $n \to \infty$. Using Fano's inequality, it can be shown that the speed of convergence for the weak secrecy criterion (5) is

$$\bar{e}_{\text{Eve}} = 1 - o(1)$$

which means that the average probability of error approaches one for increasing block length, but this can be arbitrarily slow. On the other hand, it has been recently shown in [22] that strong secrecy (6) allows for an exponential speed of convergence

$$\bar{e}_{\text{Eve}} = 1 - \mathcal{O}(2^{-\alpha n})$$

with some $\alpha > 0$. Thus, the average probability of decoding error at an eavesdropper tends to one exponentially fast for any decoder an eavesdropper may use. This demonstrates the advantage of the strong secrecy criterion compared to the weak secrecy criterion and, most importantly, it establishes a desirable and practically relevant operational meaning. Accordingly, for all results discussed in the remainder of this paper we consider the strong secrecy criterion (unless explicitly stated otherwise).

The main goal is now to determine the secrecy capacity which is the maximal achievable rate for which a code of Definition 2 can be found that ensures reliability (3) (or (4) respectively) and security (5) (or (6) respectively). This is discussed next.

D. Secrecy Capacity

The wiretap channel for perfect CSI is well studied under several aspects and its secrecy capacity can be found for instance in [2], [3], [13] and [14].

Wyner was the first to study the wiretap channel in [2]. He established the secrecy capacity for the special case of degraded channels, for which the Markov chain relation $X - Y - Z$ holds. For such channels, the eavesdropper output is always "noisier" than the output at the legitimate receiver.

Theorem 1 [2]: The secrecy capacity $C(\mathrm{W})$ of the degraded wiretap channel W is

$$C(\mathrm{W}) = \max_{X-Y-Z} \left(I(X;Y) - I(X;Z) \right)$$

with the random variables satisfying the Markov chain condition $X - Y - Z$.

The crucial idea to achieve the secrecy capacity is the following: Not all the available resources are used for the message transmission but some of them are spent for additional randomization to prevent the eavesdropper from getting any meaningful information. In more detail, for each confidential message the sender wants to transmit, there are multiple valid codewords and the stochastic encoder (1) chooses one of them uniformly at random. Now, the key insight is to choose for each message roughly $2^{nI(X;Z)}$ codewords, i.e., according to the channel quality to the eavesdropper. Thus, the eavesdropper will be saturated with useless information leaving no remaining resources for decoding the confidential message [55]. As the channel quality of the legitimate receiver roughly allows for reliable transmission at rate $I(X;Y)$, the remaining rate available for the confidential message is roughly $I(X;Y) - I(X;Z)$ as the legitimate decoder usually has to decode both: the confidential message but also the useless randomization part.

Subsequently, Csiszár and Körner have extended this formalism to broader classes of channels in [3]: Theorem 1

also holds for less noisy and more capable wiretap channels (cf. for example [56] or [57] for a discussion on less noisy and more capable channels). All these classes have in common that the legitimate channel is "stronger" than the eavesdropper channel. They further solved the general case in which there need not be an ordering between the legitimate and eavesdropper channel. In this general case, the secrecy capacity becomes the following.

Theorem 2 [3]: The secrecy capacity $C(\mathrm{W})$ of the wiretap channel W is

$$C(\mathrm{W}) = \max_{U-X-(Y,Z)} \left(I(U;Y) - I(U;Z)\right) \tag{7}$$

with the random variables satisfying the Markov chain condition $U - X - (Y, Z)$.

Compared to the degraded case in Theorem 1, there is now the need for an auxiliary random variable U. It realizes an additional randomization which is also known as *channel prefixing* as it basically creates new channels $\tilde{W} : \mathcal{U} \to \mathcal{P}(\mathcal{Y})$ and $\tilde{V} : \mathcal{U} \to \mathcal{P}(\mathcal{Z})$ from U (instead of X) to Y and Z. Now, the coding strategy of Theorem 1 is applied to these prefixed channels and for the general case, this can lead to an increase in capacity. At first, adding an additional channel sounds counter-intuitive as this can only decrease the channel qualities to the legitimate receiver and the eavesdropper, i.e., $I(U;Y) \leq I(X;Y)$ and $I(U;Z) \leq I(X;Z)$. Now, the crucial idea of channel prefixing is to find a U such that the eavesdropper channel quality is much more decreased than the legitimate channel quality. Since the resulting rate is given by the difference of both channel qualities, this would actually yield an increase in capacity, i.e., $I(U;Y) - I(U;Z) \geq I(X;Y) - I(X;Z)$. However, for wiretap channels with a "stronger" legitimate channel, there is not such an U and channel prefixing does not increase capacity. Thus, the choice $U = X$ is capacity-achieving and Theorem 1 results.

III. COMPOUND WIRETAP CHANNEL

In this section, we discuss the first scenario with a more realistic CSI assumption which is given by the concept of compound channels [19], [20]. Of interest is then the corresponding *compound wiretap channel* which models the uncertainty scenario, in which the actual channel realization is unknown to the transmitter and the legitimate receiver. It is only known to them that the actual channel realization remains constant during the entire transmission of a codeword and belongs to a known uncertainty set. This captures realistic communication conditions in which CSI is only imperfectly available at the users; for example due to inaccurate channel estimation or limited feedback schemes. Moreover, this uncertainty in CSI can also originate from active adversaries who are able to influence or control the channel state as indicated in Fig. 2.

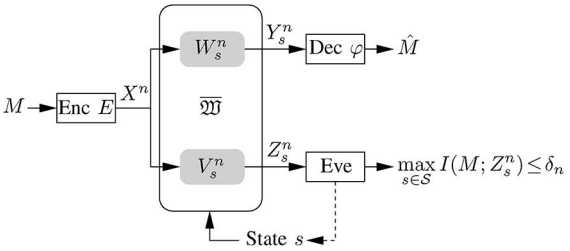

Fig. 2. *Compound wiretap channel $\overline{\mathfrak{W}}$. The actual channel realization (W_s^n, V_s^n) is unknown to the transmitter and legitimate receiver. It is only known that the corresponding state s belongs to a known uncertainty set $\mathcal{S}$. Accordingly, the eavesdropper has to be kept ignorant of M for all possible $s \in \mathcal{S}$ so that the security condition becomes $\max_{s\in\mathcal{S}} I(M;Z_s^n) \leq \delta_n$.*

A. System Model

To model the uncertainty in CSI, we introduce a state set $\mathcal{S}$ and for each channel realization $s \in \mathcal{S}$ we then have stochastic matrices $W_s : \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ and $V_s : \mathcal{X} \to \mathcal{P}(\mathcal{Z})$ describing the potential channels to the legitimate receiver and the eavesdropper, respectively. We assume discrete memoryless channels so that the probability law for transmission of block length n for channel realization $s \in \mathcal{S}$ is $W_s^n(y^n|x^n) = \prod_{i=1}^n W_s(y_i|x_i)$ and $V_s^n(z^n|x^n) = \prod_{i=1}^n V_s(z_i|x_i)$ with input and output sequences $x^n \in \mathcal{X}^n$, $y^n \in \mathcal{Y}^n$, and $z^n \in \mathcal{Z}^n$, respectively. The marginal compound channels are now given by the families of all possible channel realizations and we define

$$\overline{\mathcal{W}} = \{W_s : s \in \mathcal{S}\} \quad \text{and} \quad \overline{\mathcal{V}} = \{V_s : s \in \mathcal{S}\}.$$

Definition 3: The discrete memoryless *compound wiretap channel* $\overline{\mathfrak{W}}$ is given by the families of marginal compound channels with common input as

$$\overline{\mathfrak{W}} = \{\overline{\mathcal{W}}, \overline{\mathcal{V}}\}.$$

The actual channel realization $s \in \mathcal{S}$ which governs the transmission is unknown to the transmitter and legitimate receiver and, furthermore, there is no prior distribution on $\mathcal{S}$ assumed. Therefore, a universal strategy is needed that works for all possible channel realizations $s \in \mathcal{S}$ simultaneously, i.e., the encoder and decoder must be independent of the channel realization $s \in \mathcal{S}$. This means we seek a code (in the sense of Definition 2) that yields a small average (or maximum) probabilities of error as in (3) [or (4)] and a small information leakage as in (6) for all $s \in \mathcal{S}$

simultaneously. Accordingly, we define the average and maximum probability of error as

$$\bar{e} = \max_{s\in\mathcal{S}} \frac{1}{|\mathcal{M}|} \sum_{m\in\mathcal{M}} \sum_{x^n\in\mathcal{X}^n} \sum_{y^n:\varphi(y^n)\neq m} W_s^n(y^n|x^n)E(x^n|m)$$

and

$$e_{\max} = \max_{s\in\mathcal{S}} \max_{m\in\mathcal{M}} \sum_{x^n\in\mathcal{X}^n} \sum_{y^n:\varphi(y^n)\neq m} W_s^n(y^n|x^n)E(x^n|m).$$

The strong secrecy criterion becomes

$$\max_{s\in\mathcal{S}} I\left(M;Z_s^n\right) \leq \delta_n \tag{8}$$

with $\delta_n \to 0$ as $n \to \infty$. Taking the maximum over all possible channel realizations ensures that the legitimate users are prepared for the worst and that reliability and security is guaranteed for all possible $s \in \mathcal{S}$.

Now, the aim is to characterize the secrecy capacity of the compound wiretap channel, which is the maximal achievable rate for which a code of Definition 2 can be found such that reliability and secrecy as defined above are guaranteed.

B. No Channel State Information

We start with the scenario in which neither the transmitter nor the legitimate receiver knows the exact channel realization that governs the transmission as shown in Fig. 2. This situation has been studied for discrete memoryless channels in [21], [22] and [25].

Since for the compound wiretap channel the legitimate users do not know the actual channel and they must prepare for all possible realizations, the secrecy capacity of the compound wiretap channel cannot exceed the secrecy capacity of any wiretap channel in this family. Unfortunately, the minimum of all these secrecy capacities (worst-case secrecy capacity) is usually not a tight upper bound. The reason is that different wiretap channels may have different capacity-achieving input distributions so that the encoder has to adapt its coding strategy accordingly to achieve the worst-case secrecy capacity. However for the compound wiretap channel, the encoder must be universal and has to choose one input distribution that balances the rates for all the potential channels in the best possible way. This makes the secrecy capacity of the compound wiretap channel usually strictly smaller than its worst-case capacity. However, it immediately yields an upper bound on the compound secrecy capacity.

Proposition 1 [21]: The strong secrecy capacity of the compound wiretap channel $\overline{\mathfrak{W}}$ is upper bounded by its worst-case secrecy capacity

$$C_{\text{noCSI}}(\overline{\mathfrak{W}}) \leq \min_{s\in\mathcal{S}} \max_{U_s - X_s - (Y_s,Z_s)} \left(I(U_s;Y_s) - I(U_s;Z_s)\right)$$

for random variables $U_s - X_s - (Y_s, Z_s)$. Here, the subscripts in U_s and X_s indicate that the channel input and the channel prefixing depend on the actual channel realization $s \in \mathcal{S}$.

Applying the coding ideas for the perfect CSI case to the compound scenario at hand, yields an achievable secrecy rate as given in the following. Now, as the actual channel realization to the eavesdropper is unknown to the transmitter, it has to prepare for the worst (which is the best eavesdropper channel in this case) and chooses the randomization rate roughly as $\max_{s\in\mathcal{S}} I(U;Z_s)$, i.e., according to the best possible channel quality to the eavesdropper. This ensures that the eavesdropper will be saturated with sufficient useless information regardless of its actual channel quality. Similarly, to be on the safe side for reliable communication, the transmission rate is limited to $\min_{s\in\mathcal{S}} I(U;Y_s)$, i.e., according to the worst possible channel quality to the legitimate receiver. The rate for the confidential message follows accordingly. The auxiliary random variable U plays the role of additional channel prefixing similarly as in the case of perfect CSI in Section II.

Theorem 3 [21], [22]: An achievable strong secrecy rate for the compound wiretap channel $\overline{\mathfrak{W}}$ is

$$C_{\text{noCSI}}(\overline{\mathfrak{W}}) \geq \max_{U-X-(Y_s,Z_s)} \left(\min_{s\in\mathcal{S}} I(U;Y_s) - \max_{s\in\mathcal{S}} I(U;Z_s) \right) \tag{9}$$

for random variables $U - X - (Y_s, Z_s)$. Here, U and X are independent of the actual realization $s \in \mathcal{S}$ indicating that channel input and channel prefixing are chosen universally.

Unfortunately, a single-letter characterization of the compound secrecy capacity remains open. The secrecy capacity has only been established for the case of degraded channels, for which each possible eavesdropper channel must be degraded with respect to all possible legitimate receiver channels. This means, having two uncertainty sets, a set $\mathcal{S}$ for the legitimate receiver and a set $\mathcal{T}$ for the eavesdropper, the Markov chain relationship $X - Y_s - Z_t$ must hold for all $s \in \mathcal{S}$ and $t \in \mathcal{T}$. For this special case, the capacity has been found in [21] and [22] showing that the achievability result given before in Theorem 3 is actually

tight (with no channel prefixing, i.e., with the choice $U = X$).

Theorem 4 [21], [22]: The strong secrecy capacity of the degraded compound wiretap channel $\overline{\mathfrak{W}}$ is

$$C_{\text{noCSI}}(\overline{\mathfrak{W}}) = \max_{X - Y_s - Z_t} \left(\min_{s \in \mathcal{S}} I(X; Y_s) - \max_{t \in \mathcal{T}} I(X; Z_t) \right) \quad (10)$$

for random variables $X - Y_s - Z_t$.

These results show that, in principle, secure communication over compound wiretap channels behaves in a similar way as for the case of perfect CSI: The corresponding secrecy capacities display the same kind of structure. They further reveal how the uncertainty in the CSI affects the performance of secure communication. In particular, by comparing the compound case in (9) or (10) with the case of perfect CSI in (7), it can bee seen that the uncertainty reduces the secrecy rate in two different ways according to the two performance criteria of reliability and security: First, the uncertainty affects the reliable communication by reducing the transmission rate to the minimum of all potential channels to the legitimate receiver (resulting in the term $\min_{s \in \mathcal{S}} I(X; Y_s)$), since reliable communication must be enabled for all possible legitimate channel realizations. And second, to ensure the security of the transmitted message, the penalty in rate is increased to the maximum of all potential channels to the eavesdropper (resulting in $\max_{s \in \mathcal{S}} I(X; Z_s)$), since security must be guaranteed for all possible eavesdropper channel realizations. The assumption of simultaneously having the worst channel to the legitimate receiver and the best channel to the eavesdropper might be very pessimistic, but this is indispensable for guaranteeing reliability and security regardless of the actual channel realization.

Despite these efforts, a single-letter characterization of the secrecy capacity for the general case remains unknown until now (if it exists at all). Only a multi-letter upper bound is known which leads to a multi-letter description of the secrecy capacity.

Theorem 5 [22]: A multi-letter description of the strong secrecy capacity of the compound wiretap channel $\overline{\mathfrak{W}}$ is

$$C_{\text{noCSI}}(\overline{\mathfrak{W}}) = \lim_{n \to \infty} \frac{1}{n} \max_{U - X^n - \left(Y_s^n, Z_s^n\right)} \left(\min_{s \in \mathcal{S}} I\left(U; Y_s^n\right) - \max_{s \in \mathcal{S}} I\left(U; Z_s^n\right) \right)$$

for random variables $U - X^n - (Y_s^n, Z_s^n)$.

Thus, in principle a characterization of the secrecy capacity of the general compound wiretap channel is known in terms of entropic quantities. Although the result is given in a multi-letter form, the achievability follows by applying the single-letter result in Theorem 3 to the n-fold channels $\widetilde{W} : \mathcal{U} \to \mathcal{P}(\mathcal{Y}^n)$ and $\widetilde{V} : \mathcal{U} \to \mathcal{P}(\mathcal{Z}^n)$ which collect all the n channel uses in a single "block."

Unfortunately, the capacity result is given in a multi-letter form depending on the block length n (which tends to infinity), which makes this expression intractable to evaluate. However, such a description is still useful as it gives valuable insights and allows us to learn certain properties of the secrecy capacity as discussed later.

C. Channel State Information at Transmitter

Interestingly, the secrecy capacity of the compound wiretap channel does not increase if the legitimate receiver is informed about the actual channel realization $s \in \mathcal{S}$. An intuitive explanation of why CSI at the legitimate receiver does not improve capacity for compound channels is discussed in [58]: Even if the actual channel to the legitimate receiver is unknown, it can be estimated with arbitrary accuracy at the receiver. Then for sufficiently large block length n, the portion used for the estimation is a negligible part of n and approaches zero as $n \to \infty$.

On the other hand, an informed transmitter usually leads to an increase in secrecy capacity which is discussed next. Accordingly, we assume that we have channel state information at the transmitter (CSIT) but not at the legitimate receiver. This allows the transmitter to adapt the encoder according to the particular $s \in \mathcal{S}$ which governs the transmission. In contrast to the previous case with no CSI at all users, a single-letter characterization of the secrecy capacity has been found for this case in [22].

Theorem 6 [22]: The strong secrecy capacity of the compound wiretap channel $\overline{\mathfrak{W}}$ with CSIT is

$$C_{\text{CSIT}}(\overline{\mathfrak{W}}) = \min_{s \in \mathcal{S}} \max_{U_s - X_s - (Y_s, Z_s)} \left(I(U_s; Y_s) - I(U_s; Z_s) \right)$$

for random variables $U_s - X_s - (Y_s, Z_s)$.

Comparing the case of CSIT with the outer bound in Proposition 1 shows that the secrecy capacity of the compound wiretap channel with CSIT is at least as good that with no CSI. Moreover, it reveals that the secrecy capacity of the compound wiretap channel with CSIT actually equals the worst-case capacity. Intuitively this makes sense as in the CSIT case, the encoder can adapt its input distribution according to the actual channel. Thus, for each possible channel realization it can choose the maximizing input distribution so that the worst-case capacity becomes achievable.

Having no CSI at the transmitter forces the encoder to choose the randomization part of rate $\max_{s\in\mathcal{S}} I(X;Z_s)$ to ensure that the eavesdropper is sufficiently saturated regardless of the actual channel realization. CSIT now allows the encoder to adapt the rate of the randomization part to what is needed for the actual channel realization. A consequence is that the sizes of the codebooks differ for different channel states. This is where the potential increase in secrecy capacity arises. Interestingly, this prevents the legitimate receiver (who is not aware of the actual channel realization) to use the classical approach of decoding the confidential message and the randomization index. This necessitates a more sophisticated decoding strategy which solely decodes the confidential message but not the randomization index [22].

D. Continuity and Robustness

Next, we want to address the question of continuity and robustness which is driven by the following observation: Obviously, the secrecy capacity of the compound wiretap channel should depend in a continuous way on the underlying uncertainty set. This is because small variations in the uncertainty set should result in small variations of the secrecy capacity only. Such a continuous dependency is particularly desirable from an adversarial point of view, where the uncertainty set reflects the strategy space of malevolent adversaries.

For this purpose, we need a concept to measure the distance between two compound wiretap channels. First, we define the distance between two channels $W_1, W_2 : \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ based on the total variation distance as

$$d(W_1, W_2) = \max_{x\in\mathcal{X}} \sum_{y\in\mathcal{Y}} |W_1(y|x) - W_2(y|x)|. \tag{11}$$

Then the distance $D(\overline{\mathfrak{W}}_1, \overline{\mathfrak{W}}_2)$ between two compound wiretap channels $\overline{\mathfrak{W}}_1$ and $\overline{\mathfrak{W}}_2$ is given by the largest distance in (11) for all possible channel realizations for the legitimate and eavesdropper channel.

Theorem 7 [39]: Let $\epsilon \in (0,1)$ be arbitrary. Let $\overline{\mathfrak{W}}_1$ and $\overline{\mathfrak{W}}_2$ be two compound wiretap channels. If

$$D(\overline{\mathfrak{W}}_1, \overline{\mathfrak{W}}_2) < \epsilon,$$

then it holds that

$$\left|C(\overline{\mathfrak{W}}_1) - C(\overline{\mathfrak{W}}_2)\right| \leq \delta(\epsilon, |\mathcal{Y}|, |\mathcal{Z}|) \tag{12}$$

with $\delta(\epsilon, |\mathcal{Y}|, |\mathcal{Z}|)$ a constant depending only on the distance ϵ and the output alphabet sizes $|\mathcal{Y}|$ and $|\mathcal{Z}|$.

This result shows that the strong secrecy capacity of the compound wiretap channel is a continuous function of the uncertainty set. Thus, small variations in the uncertainty set result in only small variations of the corresponding secrecy capacity. In addition, (12) explicitly quantifies by $\delta(\epsilon, |\mathcal{Y}|, |\mathcal{Z}|)$ how much the secrecy capacity can differ when the uncertainty set of a compound wiretap channel changes by ϵ.

Basically, Theorem 7 ensures the following: If there is a "good" (i.e., capacity-achieving) code for the compound wiretap channel $\overline{\mathfrak{W}}_1$, then there exists another "good" code that achieves a similar rate over another compound wiretap channel $\overline{\mathfrak{W}}_2$ provided that they are close, i.e., $D(\overline{\mathfrak{W}}_1, \overline{\mathfrak{W}}_2) < \epsilon$. However, it does not guarantee that one particular code will possess this property. Nevertheless, it is a necessary property for bringing such concepts into practice. Since if the secrecy capacity would be discontinuous, it would be hopeless to even aim for actual codes that possess such desirable properties.

A code is called *robust* if its decoding performance at the legitimate receiver and secrecy at the eavesdropper (and therewith also the transmission rate) depend continuously on the underlying uncertainty set. We are now interested in studying whether actual codes are themselves robust against small variations in the uncertainty set. The robustness of the reliability to the legitimate receiver follows from a discussion for the classical compound channel in [19]. There it is shown that a "good" code in the sense of having small probability of decoding error, performs well also for compound channels in a certain neighborhood. This is exactly what we need. The next result shows further that the weak secrecy criterion is robust as well against small changes in the uncertainty set.

Theorem 8 [39]: Let $\overline{\mathcal{V}}_1$ be a compound channel to the eavesdropper with state set $\mathcal{S}_1$. Then for any code that achieves weak secrecy

$$\max_{s_1\in\mathcal{S}_1} \frac{1}{n} I(M; Z_{s_1}^n) = \delta_n$$

it holds that for all compound channels $\overline{\mathcal{V}}_2$ with finite state set $\mathcal{S}_2$ and $D(\overline{\mathcal{V}}_1, \overline{\mathcal{V}}_2) < \epsilon$ that

$$\max_{s_2\in\mathcal{S}_2} \frac{1}{n} I\left(M; Z_{s_2}^n\right) < \delta_n + \delta_2(\epsilon, |\mathcal{Z}|) \tag{13}$$

with $\delta_2(\epsilon, |\mathcal{Z}|)$ a constant depending only on the distance ϵ and the output alphabet size $|\mathcal{Z}|$.

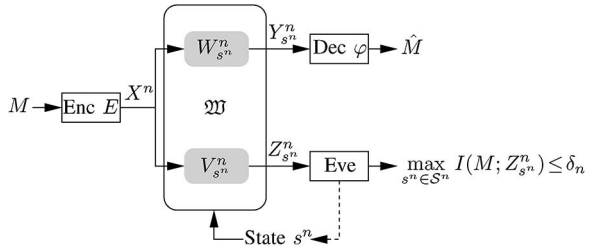

Fig. 3. ***Arbitrarily varying wiretap channel $\mathfrak{W}$. In contrast to the compound wiretap channel, the actual channel is now governed by an unknown state sequence $s^n \in \mathcal{S}^n$ of length n, which may vary in an arbitrary manner from channel use to channel use.***

This has a practically relevant consequence: A code with small information leakage rate over the eavesdropper compound channel $\overline{\mathcal{V}}_1$ has also small information leakage rate for all compound channels $\overline{\mathcal{V}}_2$ which are close, i.e., $D(\overline{\mathcal{V}}_1, \overline{\mathcal{V}}_2) < \epsilon$. In addition, the change in information leakage is explicitly quantified and bounded by (13).

Finally, it is noteworthy that these properties have been established without having a single-letter description of the secrecy capacity available. This shows that although a multi-letter characterization of the secrecy capacity as given in Theorem 5 is not efficiently computable, it is extremely useful for obtaining certain properties such as continuity or robustness.

IV. ARBITRARILY VARYING WIRETAP CHANNEL

In this section, we continue our analysis by considering secure communication over AVCs [30]–[32]. Of interest is then the corresponding *arbitrarily varying wiretap channel (AVWC)*. In contrast to the compound channel, the unknown channel realization may now vary in an unknown and arbitrary manner from channel use to channel use. This includes channel conditions such as fast fading but also captures scenarios with active adversaries, who maliciously jam the legitimate transmission as depicted in Fig. 3.

A. System Model

Similar to the compound wiretap channel, we model the uncertainty in CSI with the help of a finite state set $\mathcal{S}$. For a fixed state sequence $s^n \in \mathcal{S}^n$ of length n, the discrete memoryless channel to the legitimate receiver is now given by $W^n_{s^n}(y^n|x^n) = W^n(y^n|x^n, s^n) = \prod_{i=1}^n W(y_i|x_i, s_i)$.

Then the (marginal) AVC $\mathcal{W}$ to the legitimate receiver is defined as the family of channels for all $s^n \in \mathcal{S}^n$ as

$$\mathcal{W} = \left\{W^n_{s^n} : s^n \in \mathcal{S}^n\right\}.$$

Further, for any probability distribution $q \in \mathcal{P}(\mathcal{S})$ the *averaged channel* is defined as

$$W_q(y|x) = \sum_{s \in \mathcal{S}} W(y|x, s)q(s)$$

for all $x \in \mathcal{X}$ and $y \in \mathcal{Y}$.

We can define the channel to the eavesdropper in a similar way. For given state sequence $s^n \in \mathcal{S}^n$ the discrete memoryless channel is given as $V^n_{s^n}(z^n|x^n) = V^n(z^n|x^n, s^n) = \prod_{i=1}^n (z_i|x_i, s_i)$. We also set $\mathcal{V} = \{V^n_{s^n} : s^n \in \mathcal{S}^n\}$ and $V_q(z|x) = \sum_{s \in \mathcal{S}} V(z|x, s)q(s)$ for $q \in \mathcal{P}(\mathcal{S})$.

Definition 4: The discrete memoryless *arbitrarily varying wiretap channel (AVWC)* $\mathfrak{W}$ is given by the families of marginal AVCs with common input as

$$\mathfrak{W} = \{\mathcal{W}, \mathcal{V}\}.$$

B. Code Concepts

For communication over AVCs it makes a critical difference whether unassisted or common randomness (CR) assisted codes are used. The unassisted capacity can be zero while CR-assisted codes allow for communication at a non-zero rate [30]–[32]. In the following we introduce these code concepts for the AVWC.

1) Unassisted Codes: As we will deal with different code concepts in the following, we refer to (n, M_n)-codes $\mathcal{C}$ of Definition 2 as *unassisted* codes to distinguish them from more sophisticated code concepts. The term "unassisted" refers to the fact that encoder (1) and decoder (2) are chosen and fixed prior to the message transmission (i.e., they have to be universal and their choice cannot be coordinated in any way).

Now, the reliability and security requirements have to take all state sequences $s^n \in \mathcal{S}^n$ of length n into account. Accordingly, the probability of error for message $m \in \mathcal{M}$ and state sequence $s^n \in \mathcal{S}^n$ is given by

$$e(m, s^n) = \sum_{x^n \in \mathcal{X}^n} \sum_{y^n : \varphi(y^n) \neq m} W^n_{s^n}(y^n|x^n)E(x^n|m).$$

Since reliability must be guaranteed for all $s^n \in \mathcal{S}^n$, the average probability of error is

$$\bar{e} = \max_{s^n \in \mathcal{S}^n} \frac{1}{|\mathcal{M}|} \sum_{m \in \mathcal{M}} e(m, s^n). \tag{14}$$

Similarly the strong secrecy criterion becomes

$$\max_{s^n \in \mathcal{S}^n} I\left(M; Z^n_{s^n}\right) \leq \delta_n \tag{15}$$

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

Then the *unassisted strong secrecy capacity* $C(\mathfrak{W})$ of the AVWC $\mathfrak{W}$ is given by the maximal achievable rate for which an unassisted code of Definition 2 can found that ensures reliability (14) and secrecy (15).

Note that we only consider the average probability of error criterion in the following (and not the maximum probability as well as previously done). While for perfect CSI and compound channels the secrecy capacity turns out to be the same for both average and maximum error, the situation completely changes for AVCs. Already for the single-user case without secrecy, the capacity under average and maximum error differs and the capacity under maximum error is still unknown. Furthermore, this includes the famous zero-error problem of Shannon as a special case [59].

When the AVWC is studied from an adversarial point of view, the state sequence does not originate solely from channel uncertainty but is controlled by a malevolent adversary [35]. Then the reliability and security conditions already indicate that there are different strategies possible for the adversary, since the particular state sequence influences the decoding performance at the legitimate receiver (14) and, at the same time, has an impact on the information leakage to the eavesdropper (15). Accordingly, one approach would be to disturb the legitimate communication as much as possible by choosing the state sequence in such a way that the probability of error (14) becomes as large as possible. Contrary to that, the adversary can choose the state sequence to maximize the information leakage (15). Of course, any strategy in between is also a valid jamming strategy, and thus, the legitimate users must be prepared for all possible strategies which is reflected by the maximum in (14) and (15).

Unfortunately, such unassisted approaches work only for certain channel configurations. In particular, if the AVC possesses the so-called property of symmetrizability, then unassisted codes will yield a zero capacity as discussed next.

Definition 5: An AVC $\mathcal{W}$ is called *symmetrizable* if there exists a channel (stochastic matrix) $\sigma : \mathcal{X} \to \mathcal{P}(\mathcal{S})$ such that

$$\sum_{s \in \mathcal{S}} W(y|x,s)\sigma(s|x') = \sum_{s \in \mathcal{S}} W(y|x',s)\sigma(s|x) \tag{16}$$

holds for all $x, x' \in \mathcal{X}$ and $y \in \mathcal{Y}$.

Writing the left hand side of (16) as $\widetilde{W}(y|x,x') = \sum_{s \in \mathcal{S}} W(y|x,s)\sigma(s|x')$ reveals the following meaning of symmetrizability: The resulting channel $\widetilde{W}$ is symmetric in both inputs x and x' so that

$$\widetilde{W}(y|x,x') = \widetilde{W}(y|x',x).$$

Thus, roughly speaking, a symmetrizable AVC can "simulate" a valid channel input making it impossible for the decoder to decide on whether x was sent and x' is the interference or if it is the other way and x' was sent and x is the interference. That this actually leads to a zero unassisted capacity is further elaborated in the following.

Let $x^n_m \in \mathcal{X}^n$ and $m \in \mathcal{M}$ be arbitrary codewords. Following the interpretation that for a symmetrizable AVC the interfering sequences can look like valid channel inputs, we set $s^n_m = x^n_m$ for all $m \in \mathcal{M}$. For the expected probability of error, we obtain for each pair of codewords $(k,l) \in \mathcal{M} \times \mathcal{M}$ with $k \neq l$ the following:

$$\begin{aligned}
&\mathbb{E}\left[e\left(k, S^n_l\right)\right] + \mathbb{E}\left[e\left(l, S^n_k\right)\right] \\
&\quad = \mathbb{E}\left[W^n\left(\left(\varphi^{-1}(k)\right)^c | x^n_k, S^n_l\right)\right] + \mathbb{E}\left[W^n\left(\left(\varphi^{-1}(l)\right)^c | x^n_l, S^n_k\right)\right] \\
&\quad = \mathbb{E}\left[W^n\left(\left(\varphi^{-1}(k)\right)^c | x^n_k, S^n_l\right)\right] + \mathbb{E}\left[W^n\left(\left(\varphi^{-1}(l)\right)^c | x^n_k, S^n_l\right)\right] \\
&\quad \geq \mathbb{E}\left[W^n\left(\left(\varphi^{-1}(k)\right)^c | x^n_k, S^n_l\right)\right] + \mathbb{E}\left[W^n\left(\varphi^{-1}(k) | x^n_k, S^n_l\right)\right] \\
&\quad = \mathbb{E}\left[W^n\left(\left(\varphi^{-1}(k)\right)^c \cup \varphi^{-1}(k) | x^n_k, S^n_l\right)\right] \\
&\quad = 1
\end{aligned}$$

where the first equality follows from the definition of probability of error and the second equality from the fact that the AVC $\mathcal{W}$ is symmetrizable, cf. Definition 5. This means that for each pair of codewords the expectation of the probabilities of error is lower bounded by one. Then averaging over all codewords leads to

$$\frac{1}{|\mathcal{M}|} \sum_{l=1}^{M_n} \mathbb{E}\left[\bar{e}\left(S^n_l\right)\right] \geq \frac{1}{4}$$

which implies that $\mathbb{E}[\bar{e}(S^n_l)] \geq 1/4$ holds for at least one $l \in \mathcal{M}$. But if the average probability of error is bounded from below by a positive constant, reliable communication is not possible so that the unassisted capacity is zero in this case [32], [34], [35]. This necessitates the use of more sophisticated strategies based on *common randomness (CR)* as discussed next.

2) CR-Assisted Codes: CR can be realized by some common satellite signal or can be obtained by common synchronization procedures. It is modeled by a random variable Γ which takes values in a finite set $\mathcal{G}_n$ according to a distribution $P_\Gamma \in \mathcal{P}(\mathcal{G}_n)$. This enables the transmitter

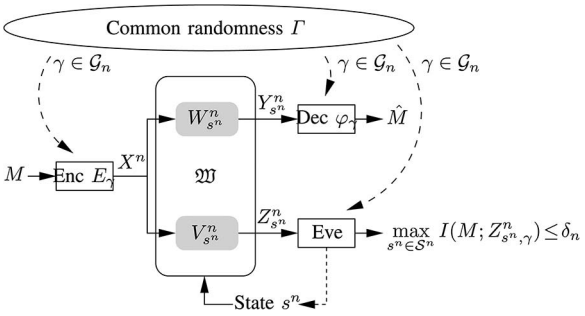

Fig. 4. *All users (including the eavesdropper) have access to a common random source and can adapt their encoder and decoders accordingly to the actual CR realization $\gamma \in \mathcal{G}_n$.*

and legitimate receiver to coordinate their choice of encoder (1) and decoder (2) according to the realization $\gamma \in \mathcal{G}_n$. This scenario is depicted in Fig. 4.

Definition 6: A *CR-assisted* $(n, M_n, \mathcal{G}_n, P_\Gamma)$*-code* $\mathcal{C}_{\mathrm{CR}}$ is given by a family of unassisted codes

$$\{\mathcal{C}(\gamma) : \gamma \in \mathcal{G}_n\}$$

together with a random variable Γ taking values in $\mathcal{G}_n$ with $|\mathcal{G}_n| < \infty$ according to $P_\Gamma \in \mathcal{P}(\mathcal{G}_n)$. The rate of this code is defined as $(1/n)\log M_n$.

The natural way of extending the reliability and secrecy requirements from above to CR-assisted codes is done by further averaging over all possible CR realizations $\gamma \in \mathcal{G}_n$. Thus, the mean average probability of error is

$$\bar{e}_{\mathrm{CR}} = \max_{s^n \in \mathcal{S}^n} \frac{1}{|\mathcal{M}|} \sum_{m \in \mathcal{M}} \sum_{\gamma \in \mathcal{G}_n} \sum_{x^n \in \mathcal{X}^n} \times \sum_{y^n : \varphi(y^n) \neq m} W^n_{s^n}(y^n|x^n) E_\gamma(x^n|m) P_\Gamma(\gamma) \quad (17)$$

and the strong secrecy criterion becomes

$$\max_{s^n \in \mathcal{S}^n} \sum_{\gamma \in \mathcal{G}_n} I\left(M; Z^n_{s^n,\gamma}\right) P_\Gamma(\gamma) \leq \delta_n \quad (18)$$

with $\delta_n \to 0$ as $n \to \infty$. Here, $Z^n_{s^n,\gamma}$ denotes the channel output at the eavesdropper for state sequence $s^n \in \mathcal{S}^n$ and CR realization $\gamma \in \mathcal{G}_n$. The secrecy criterion (18) can further be strengthened by replacing the average over all CR realizations by the maximum, i.e.

$$\max_{s^n \in \mathcal{S}^n} \max_{\gamma \in \mathcal{G}_n} I\left(M; Z^n_{s^n,\gamma}\right) \leq \delta_n. \quad (19)$$

This was first considered in [38] and, surprisingly, strengthening the secrecy criterion from (18) to (19) comes at no cost in terms of secrecy capacity as we will see later. The stronger secrecy criterion has the advantage that in this case, the communication is secure even if the eavesdropper is aware of the actual CR realization $\gamma \in \mathcal{G}_n$.

In particular, the latter assumption of an eavesdropper that is aware of the CR realization is meaningful, since otherwise, the legitimate users could immediately use this resource to create a secret key corresponding to the size of the CR. Such a secret key can then be used as a *one-time pad* to keep the communication secure [60]–[64].

The *CR-assisted secrecy capacity* $C_{\mathrm{CR}}(\mathfrak{W})$ of the AVWC $\mathfrak{W}$ is given by the maximal achievable rate for which a CR-assisted code of Definition 6 can be found that ensures reliability (17) and security (18) (or (19) respectively).

C. Capacity Results

Recently, there has been some effort toward understanding the secrecy capacity of the AVWC [33]–[39], which is reviewed next.

1) Unassisted Secrecy Capacity: If no coordination resources such as CR are available to the legitimate users, unassisted codes with fixed encoder and decoders must be used. The corresponding unassisted secrecy capacity has been completely characterized in [34] and [38].

Theorem 9 [34], [38]: The unassisted strong secrecy capacity $C(\mathfrak{W})$ of the AVWC $\mathfrak{W}$ possesses the following properties:

1) if $\mathcal{W}$ is symmetrizable, then $C(\mathfrak{W}) = 0$;
2) if $\mathcal{W}$ is nonsymmetrizable, then $C(\mathfrak{W}) = C_{\mathrm{CR}}(\mathfrak{W})$.

The unassisted secrecy capacity is characterized in terms of its CR-assisted secrecy capacity and no longer by entropic quantities only (as for the perfect CSI or compound case). It displays a dichotomous behavior similar to the single-user AVC: the unassisted secrecy capacity $C(\mathfrak{W})$ either equals the CR-assisted secrecy capacity $C_{\mathrm{CR}}(\mathfrak{W})$ or else is zero. It is noteworthy that this behavior depends only on the symmetrizability of the legitimate channel and not on the channel to the eavesdropper. As a consequence, the unassisted secrecy capacity can be zero even if corresponding entropic quantities (such as the mutual information between the input and the output) are positive.

As the CR-assisted secrecy capacity is a natural upper bound on the unassisted secrecy capacity, this result further reveals the following interesting observation: When the AVC to the legitimate receiver is nonsymmetrizable, there is no gain in rate by using CR-assisted codes as the unassisted and CR-assisted secrecy capacities are equal. CR as a coordination resource will only help to enable communication in the case of symmetrizable channels as we will see next.

2) CR-Assisted Secrecy Capacity: CR allows the transmitter and receiver to coordinate their choices of encoder and decoder. This case has been studied in [33]–[37] and [39], but a single-letter characterization of the CR-assisted secrecy capacity remains unknown until now. Only a multi-letter description has been found in [37].

Theorem 10 [37]: A multi-letter description of the CR-assisted strong secrecy capacity $C_{\mathrm{CR}}(\mathfrak{W})$ of the AVWC $\mathfrak{W}$ is

$$C_{\mathrm{CR}}(\mathfrak{W}) = \lim_{n\to\infty} \frac{1}{n} \max_{U - X^n - \left(Y_q^n, Z_{s^n}^n\right)} \times \left(\min_{q\in\mathcal{P}(\mathcal{S})} I\left(U; Y_q^n\right) - \max_{s^n\in\mathcal{S}^n} I\left(U; Z_{s^n}^n\right) \right)$$

with Y_q^n the random variable associated with the output of the averaged channel $W_q^n = \sum_{s^n\in\mathcal{S}^n} q^n(s^n) W_{s^n}$, $q \in \mathcal{P}(\mathcal{S})$.

Recall that the multi-letter characterization for the compound wiretap channel in Theorem 5 relies on a single-letter achievability result of Theorem 3. In contrast to that, the multi-letter version for the AVWC in Theorem 10 above follows already from a multi-letter achievable secrecy rate. A single-letter achievable secrecy rate that works in general remains unknown. Only for the special case of a best channel to the eavesdropper, a single-letter secrecy rate is known [33], [34].

For this purpose, a channel $V_{q^*} \in \{V_q : q \in \mathcal{P}(\mathcal{S})\}$ such that all other channels from this set are degraded versions of V_{q^*} is called the best channel to the eavesdropper. This means that V_{q^*} is a best channel to the eavesdropper if

$$X - Z_{q^*} - Z$$

forms a Markov chain for all $q \in \mathcal{P}(\mathcal{S})$ with Z_q the random variable associated with the output of the averaged channel V_q, $q \in \mathcal{P}(\mathcal{S})$. If this condition is satisfied, a single-letter achievable CR-assisted secrecy rate is known and given in the following theorem.

Theorem 11 [34]: If there exists a best channel to the eavesdropper, an achievable CR-assisted strong secrecy rate for the AVWC $\mathfrak{W}$ is

$$C_{\mathrm{CR}}(\mathfrak{W}) \geq \max_{X - (Y_q, Z_q)} \left(\min_{q\in\mathcal{P}(\mathcal{S})} I(X; Y_q) - \max_{q\in\mathcal{P}(\mathcal{S})} I(X; Z_q) \right)$$

with Y_q and Z_q the random variables associated with the outputs of the averaged channels W_q and V_q, $q \in \mathcal{P}(\mathcal{S})$.

D. Stability, Continuity, and Robustness

The unassisted secrecy capacity in Theorem 9 reveals that the symmetrizability of the legitimate AVC controls whether the unassisted secrecy capacity is zero or positive. However, it does not specify the sensitivity of the AVC on the underlying uncertainty set meaning how rapidly the AVC can change from non-symmetrizable to symmetrizable. This is addressed by the next result.

Theorem 12 [38]: If the unassisted strong secrecy capacity $C(\mathfrak{W})$ of the AVWC $\mathfrak{W}$ satisfies $C(\mathfrak{W}) > 0$, then there is an $\epsilon > 0$ such that for all AVWCs $\mathfrak{W}_\epsilon$ satisfying $d(\mathfrak{W}, \mathfrak{W}_\epsilon) \leq \epsilon$ we have $C(\mathfrak{W}_\epsilon) > 0$.

This result displays the stability of positivity of the unassisted secrecy capacity: Wherever it is positive, i.e., $C(\mathfrak{W}) > 0$, it remains positive in a certain neighborhood. This means that small changes in the uncertainty set will not change the AVC from being non-symmetrizable to being symmetrizable.

We want to further explore the question of continuity for the AVWC. For this purpose, we define the function

$$F(\mathcal{W}) = \min_{\sigma:\mathcal{X}\to\mathcal{P}(\mathcal{S})} \left(\max_{x\neq x'} \sum_{y\in\mathcal{Y}} \left| \sum_{s\in\mathcal{S}} W(y|x', s)\sigma(s|x) - \sum_{s\in\mathcal{S}} W(y|x, s)\sigma(s|x') \right| \right). \quad (20)$$

This function resembles quantities from the concept of symmetrizability in Definition 5 and is a continuous function of the AVC $\mathcal{W}$ from the transmitter to the legitimate receiver. We observe that the AVC $\mathcal{W}$ is symmetrizable if and only if $F(\mathcal{W}) = 0$. With this we get a characterization of when the unassisted secrecy capacity $C(\mathfrak{W})$ displays a discontinuous behavior: The AVWC $\mathfrak{W}$ changes from non-symmetrizable to symmetrizable and the capacity breaks down to zero.

Theorem 13 [38]: The AVWC $\mathfrak{W}$ is a discontinuous point of $C(\mathfrak{W})$ if and only if the following holds:

1) $C_{\mathrm{CR}}(\mathfrak{W}) > 0$;
2) $F(\mathcal{W}) = 0$ and for every $\epsilon > 0$ there is a finite $\mathcal{W}_\epsilon$ with $d(\mathcal{W}, \mathcal{W}_\epsilon) \leq \epsilon$ and $F(\mathcal{W}_\epsilon) > 0$.

The characterization of the discontinuity in Theorem 13 depends on the CR-assisted secrecy capacity $C_{\mathrm{CR}}(\mathfrak{W})$ and the function $F(\mathcal{W})$, which are itself both continuous. Thus, interestingly, the *discontinuity* behavior of $C(\mathfrak{W})$ is completely characterized by two *continuous* functions.

When the CR-assisted secrecy capacity $C_{\mathrm{CR}}(\mathfrak{W})$ itself is zero, then the unassisted secrecy capacity $C(\mathfrak{W})$ must be zero as well and there cannot be any discontinuity. This is the motivation for the first condition in Theorem 13. The second condition relates the question of discontinuity to the function $F(\mathcal{W})$ in (20) and therewith solely to the symmetrizability of the legitimate channel. The interesting

consequence is that $C(\mathfrak{W})$ is always a continuous function of the eavesdropper channel, while the discontinuity comes from the legitimate channel only.

The first example of discontinuity of AVWCs appeared in [39]. The construction presented in that work used an AVC to the legitimate receiver that satisfies the second condition in Theorem 13. Furthermore, the eavesdropper AVC in [39] was assumed to be useless with zero capacity (i.e. each input symbol is mapped with equal probability to all possible output symbols). The result in Theorem 13 on the other hand allows one to show discontinuity of a huge class of AVWCs. Now, all eavesdropper AVCs can be considered that result in $C_{\mathrm{CR}}(\mathfrak{W}) > 0$ so that the first condition in Theorem 13 is satisfied as well.

This discussion has revealed that the unassisted secrecy capacity of the AVWC depends in a discontinuous way on the underlying uncertainty set. In particular, Theorem 13 shows that it is possible that small changes in the uncertainty set can lead to a dramatic loss in capacity: The unassisted secrecy capacity breaks down to zero once the legitimate channel becomes symmetrizable. On the other hand, the following result shows that CR stabilizes the communication in the sense that the corresponding CR-assisted secrecy capacity possesses the desired continuous behavior: Small changes in the uncertainty set result in small changes in capacity only.

Theorem 14 [38], [39]: Let $\epsilon \in (0,1)$ be arbitrary. Let $\mathfrak{W}_1$ and $\mathfrak{W}_2$ be two AVWCs. If

$$D(\mathfrak{W}_1, \mathfrak{W}_2) < \epsilon$$

then it holds that

$$|C_{\mathrm{CR}}(\mathfrak{W}_1) - C_{\mathrm{CR}}(\mathfrak{W}_2)| \leq \delta(\epsilon, |\mathcal{Y}|, |\mathcal{Z}|) \tag{21}$$

with $\delta(\epsilon, |\mathcal{Y}|, |\mathcal{Z}|)$ a constant depending only on the distance ϵ and the output alphabet sizes $|\mathcal{Y}|$ and $|\mathcal{Z}|$.

Similarly as for the compound wiretap channel in Section III-D an important question is whether this continuous behavior of the CR-assisted secrecy capacity holds for actual codes as well. The following result answers this question and shows that a "good" CR-assisted code that realizes weak secrecy performs also well for all channels in a certain neighborhood.

Theorem 15 [39]: Let $\mathcal{V}_1$ be an AVC to the eavesdropper with state set $\mathcal{S}_1$. Then for any code that achieves weak secrecy

$$\max_{s_1^n \in \mathcal{S}_1^n} \sum_{\gamma \in \mathcal{G}_n} \frac{1}{n} I\left(M; Z^n_{s_1^n, \gamma}\right) P_\Gamma(\gamma) = \delta_n$$

it holds that for all AVCs $\mathcal{V}_2$ with finite state set $\mathcal{S}_2$ and $D(\mathcal{V}_1, \mathcal{V}_2) < \epsilon$ that

$$\max_{s_2^n \in \mathcal{S}_2^n} \sum_{\gamma \in \mathcal{G}_n} \frac{1}{n} I\left(M; Z^n_{s_2^n, \gamma}\right) P_\Gamma(\gamma) < \delta_n + \delta_2(\epsilon, |\mathcal{Z}|)$$

with $\delta_2(\epsilon, |\mathcal{Z}|)$ a constant depending only on the distance ϵ and the output alphabet size $|\mathcal{Z}|$.

Thus, a CR-assisted code with small information leakage rate over the eavesdropper AVC also has small information leakage rate for all AVCs that are close. In [39] it has been further shown that the robustness of the weak secrecy criterion does not only hold for CR-assisted codes as in Theorem 15 but also for unassisted codes. This is particularly remarkable as it yields yield the following consequence: Unassisted secrecy codes are robust in the secrecy criterion and the discontinuity comes only from the legitimate link. As the secrecy criterion is already robust for unassisted codes, this further underlines the observation that CR is mainly for stabilizing the communication to the legitimate user and does not provide any gain in secrecy.

E. Super-Activation

Resource allocation is an important issue for wireless communication systems and, usually, the overall capacity of such systems is given by the sum of the capacities of the orthogonal sub-systems. Accordingly, a system consisting of two orthogonal AVCs, where both are "useless" in the sense of zero capacity, has a whole capacity of zero as well. This reflects the world view of classical additivity of basic resources in the sense that "$0 + 0 = 0$." In contrast to that, in quantum information theory there are scenarios possible which allow for *super-activation* of two orthogonal "useless" channels each with zero capacity such that the overall system has non-zero capacity, i.e., "$0 + 0 > 0$." To the best of our knowledge such phenomena of super-activation are only known in the area of quantum information theory [41], [42].

Surprisingly, in [35] it has been demonstrated for the first time that super-activation can also happen in the classical non-quantum world. In more detail, it has been demonstrated that two orthogonal AVWCs $\mathfrak{W}_1$ and $\mathfrak{W}_2$, where both are "useless" in the sense that they have zero secrecy capacity $C(\mathfrak{W}_1) = C(\mathfrak{W}_2) = 0$, can be jointly used to super-activate the whole system to allow for non-zero secrecy rates, i.e., $C(\mathfrak{W}_1 \otimes \mathfrak{W}_2) > C(\mathfrak{W}_1) + C(\mathfrak{W}_2) = 0$, where $\mathfrak{W}_1 \otimes \mathfrak{W}_2$ denotes the joint use of both orthogonal AVWCs. Joint use refers to the approach of designing the encoder and decoder jointly for both orthogonal AVCs (in contrast to the classical approach in which individual encoders and decoders are used; one independent pair for each sub-channel). Recently, this phenomenon of super-activation has been completely characterized for AVWCs in [35] and [38].

Theorem 16 [38]: Let $\mathfrak{W}_1$ and $\mathfrak{W}_2$ be two AVWCs. Then the following properties hold:

1) If $C(\mathfrak{W}_1) = C(\mathfrak{W}_2) = 0$, then

$$C(\mathfrak{W}_1 \otimes \mathfrak{W}_2) > 0$$

if and only if $\mathcal{W}_1 \otimes \mathcal{W}_2$ is non-symmetrizable and $C_{\mathrm{CR}}(\mathfrak{W}_1 \otimes \mathfrak{W}_2) > 0$. If $\mathfrak{W}_1$ and $\mathfrak{W}_2$ can be super-activated it holds that

$$C(\mathfrak{W}_1 \otimes \mathfrak{W}_2) = C_{\mathrm{CR}}(\mathfrak{W}_1 \otimes \mathfrak{W}_2).$$

2) If C_{CR} shows no super-activation for $\mathfrak{W}_1$ and $\mathfrak{W}_2$, then super-activation of C can only happen if $\mathcal{W}_1$ is non-symmetrizable and $\mathcal{W}_2$ is symmetrizable and $C_{\mathrm{CR}}(\mathfrak{W}_1) = 0$ and $C_{\mathrm{CR}}(\mathfrak{W}_2) > 0$. The statement is independent of the specific labeling.
3) There exist AVWCs that exhibit the behavior according to the second property.

Super-activation appears for example in the following channel configuration: Assume there are two orthogonal AVWCs whose unassisted secrecy capacities are zero. Thereby, one zero unassisted secrecy capacity stems from the fact that its AVC to the legitimate receiver is symmetrizable. For the other AVWC, the zero secrecy capacity comes from the fact that the eavesdropper AVC is "stronger" than the legitimate AVC. However, as this legitimate AVC supports a positive rate (although non-secure), it can be used to transmit information to the legitimate receiver (and eavesdropper) to generate CR. Then the legitimate users can use CR-assisted codes resulting in a positive secrecy capacity.

The previous result provides a complete characterization of when super-activation is possible. Super-activation is not an isolated phenomenon of orthogonal AVWCs. In fact, the following result shows that whenever super-activation is possible for two orthogonal AVWCs, it occurs for all AVWCs that are sufficiently close to them.

Theorem 17 [40]: Let $\mathfrak{W}_1$ and $\mathfrak{W}_2$ be two useless orthogonal AVWCs that can be super-activated. Then there exists an $\epsilon > 0$ such that all orthogonal AVWCs $\mathfrak{W}_1'$ and $\mathfrak{W}_2'$ that satisfy

$$D(\mathcal{W}_1, \mathcal{W}_1') < \epsilon, \quad D(\mathcal{W}_2, \mathcal{W}_2') < \epsilon$$

and

$$C_{\mathrm{CR}}(\mathfrak{W}_1' \otimes \mathfrak{W}_2') > 0$$

can be super-activated as well.

This result further reveals the following interesting behavior: In terms of super-activation, the legitimate AVC is much more important than the eavesdropper AVC. Specifically, there is no requirement on the distance of the eavesdropper channels.

Furthermore, super-activation leads to a more robust system which is continuous as shown in the following result. This is particularly noteworthy as such continuous behavior cannot be guaranteed in general for a single AVC.

Theorem 18 [40]: Let $\mathfrak{W}_1$ and $\mathfrak{W}_2$ be two useless orthogonal AVWCs that can be super-activated. Then the unassisted strong secrecy capacity $C(\mathfrak{W}_1' \otimes \mathfrak{W}_2')$ depends in a continuous way on the channels $\mathfrak{W}_1'$ and $\mathfrak{W}_2'$ with $D(\mathfrak{W}_i, \mathfrak{W}_i') < \epsilon$, $i = 1, 2$. Here, ϵ depends only on the orthogonal AVCs $\mathcal{W}_1$ and $\mathcal{W}_2$ to the legitimate receivers.

The previous discussion has shown that super-activation mostly depends on the legitimate AVC and is robust in the eavesdropper AVC. Specifically, it suffices to require the eavesdropper AVC to be in such a way that the resulting CR-assisted secrecy capacity of the corresponding AVWC is positive. Then the capacity is continuous in the eavesdropper AVC. Accordingly, it is worth studying the legitimate AVC on its own which results in the single-user AVC without secrecy requirement.

Interestingly, this problem of reliable message transmission over orthogonal AVCs is already implicitly addressed by Shannon's question of the additivity of the zero error capacity [65]. Specifically, it has been shown in [59] that the capacity of the AVC under the maximum error criterion includes the zero error capacity as a special case. Shannon conjectured the zero error capacity to be additive and it was Alon who presented a counter-example in [66] showing the capacity of reliable message transmission over orthogonal AVCs under the maximum error criterion is super-additive. This can be seen as the first contribution toward understanding the behavior of the capacity of orthogonal AVCs. In [40] the capacity of orthogonal AVCs under the average error criterion is completely characterized. It is shown that the capacity possesses the property of super-additivity. However, super-activation is not possible for public message transmission without secrecy constraints, making it a unique feature of secure communication over AVCs.

V. CONCLUSION

Information theoretic approaches to security establish reliable communication and data confidentiality jointly at the physical layer and have therefore been intensively examined as a complement to current cryptographic approaches at higher layers. Current studies rely heavily on the provision of accurate channel state information to the legitimate users. This is quite challenging especially for the channels to malevolent eavesdroppers who will not

share any information about their channels to make eavesdropping even harder. Moreover, besides passive eavesdroppers who simply eavesdrop upon the communication, there might be active adversaries who maliciously jam and influence the legitimate communication. Accordingly, there is the need to develop communication schemes that are robust against such uncertainties and adversarial attacks.

In this paper we have reviewed information theoretic concepts that model secure communication under channel uncertainty and adversarial attacks. The compound wiretap channel refers to the communication scenario with an unknown but constant channel. This model captures not only the effects of the wireless medium or practical limitations such as imperfect channel estimation or limited feedback schemes, it also describes simple adversarial attacks in which the actual channel realization is chosen by the adversary. The secrecy capacity has been established for several special cases such as degraded channels, channel state information at the transmitter, or certain MIMO configurations. However, a general single-letter formula for the secrecy capacity in this setting remains unknown; only a multi-letter description is known. Although such a multi-letter description is not efficiently computable, it provides already structural insights: the secrecy capacity of the compound wiretap channel is continuous in the uncertainty set. Thus, small variations in the uncertainty result only in small variations in the secrecy capacity.

The second model under investigation was the arbitrarily varying wiretap channel. In addition to the compound channel, the unknown channel is now allowed to vary in an unknown and arbitrary manner from channel use to channel use. This model includes further effects such as fast fading channels and malevolent adversaries who maliciously jam the legitimate communication. Preparing against such fluctuating channel conditions requires sophisticated coding strategies. Indeed, for AVWCs it has been shown that unassisted strategies may result in zero capacity while CR-assisted strategies allow for non-zero communication rates. Similar to the compound wiretap channel, there is no single-letter description of the secrecy capacity known, not even for special cases. Recently, at least a multi-letter characterization of the CR-assisted secrecy capacity in this case has been found. However, the unassisted secrecy capacity is completely characterized in terms of its CR-assisted secrecy capacity. In contrast to the compound wiretap channel, the unassisted secrecy capacity of the AVWC is not continuous in the uncertainty set. Thus, small variations in the uncertainty set or strategy space of the adversary might lead to a dramatic loss in secrecy capacity. Interestingly, this discontinuity stems only from the legitimate channel and not from the eavesdropper channel, making the secrecy criterion robust. However, the CR-assisted secrecy capacity is continuous in the uncertainty set. For secure communication over AVCs the new phenomenon of super-activation appear. Two AVWCs, each useless by itself in the sense of zero secrecy capacity, can be combined to super-activate the whole system to allow for non-zero secrecy rates.

In this paper, we have limited ourselves to discrete memoryless channels. Naturally, of practical relevance are Gaussian channels and in particular MIMO configurations. While there exist some few studies of such cases for the compound wiretap channel, there is not a single work for the AVWC to the best of our knowledge. Another limitation of this paper is that we have solely studied the wiretap channel which is the simplest scenario for secure communication involving one legitimate transmitter-receiver pair and one eavesdropper. Of practical relevance are further multi-user scenarios. While for the broadcast channel with confidential messages or the multiple access wiretap channel, there are some first results, this progress is far from being exhaustive. ■

Acknowledgment

H. Boche would like to thank Dr. R. Plaga, Federal Office for Information Security (BSI), for motivating and fruitful discussions that lead to these results. Results for the compound wiretap channel were presented at the industrial board meeting on "Information Security" of the German Ministry of Education and Research (BMBF) in Bonn, Germany, May 2014 and results for the AVWC were presented at the industrial board meeting on "Information Security" of the German Ministry of Education and Research (BMBF) in Bonn, Germany, May 2015.

REFERENCES

[1] C. E. Shannon, "Communication theory of secrecy systems," *Bell Syst. Tech. J.*, vol. 28, no. 4, pp. 656–715, Oct. 1949.

[2] A. D. Wyner, "The wire-tap channel," *Bell Syst. Tech. J.*, vol. 54, pp. 1355–1387, Oct. 1975.

[3] I. Csiszár and J. Körner, "Broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. 24, no. 3, pp. 339–348, May 1978.

[4] Y. Liang, H. V. Poor, and S. Shamai (Shitz), "Information theoretic security," *Found. Trends Commun. Inf. Theory*, vol. 5, no. 4/5, pp. 355–580, 2009.

[5] R. Liu and W. Trappe, Eds., *Securing Wireless Communications at the Physical Layer*. New York, NY, USA: Springer-Verlag, 2010.

[6] E. A. Jorswieck, A. Wolf, and S. Gerbracht, "Secrecy on the physical layer in wireless networks," *Trends Telecommun. Tech.*, pp. 413–435, Mar. 2010.

[7] M. Bloch and J. Barros, *Physical-Layer Security: From Information Theory to Security Engineering*. Cambridge, U.K.: Cambridge Univ. Press, 2011.

[8] X. Zhou, L. Song, and Y. Zhang, Eds., *Physical Layer Security in Wireless Communications*. Boca Raton, FL, USA: CRC Press, 2013.

[9] R. F. Schaefer and H. Boche, "Physical layer service integration in wireless networks—Signal processing challenges," *IEEE Signal Process. Mag.*, vol. 31, no. 3, pp. 147–156, May 2014.

[10] Deutsche Telekom AG Laboratories "Next generation mobile networks: (R)evolution in mobile communications," *Technology Radar Edition III/2010, Feature Paper*, 2010. [Online]. Available: http://www.lti.ei.tum.de/index.php?id=boche

[11] U. Helmbrecht and R. Plaga, "New challenges for IT-security research in ICT," in *World Federation of Scientists International Seminars on Planetary Emergencies*, Erice, Italy, Aug. 2008, pp. 1–6.

[12] G. Fettweis *et al.*, "The Tactile Internet," ITU-T Tech. Watch Rep., Tech. Rep., Aug. 2014. [Online]. Available: http://www.itu.int/oth/T2301000023/en

[13] I. Csiszár, "Almost independence and secrecy capacity," *Probl. Pered. Inform.*, vol. 32, no. 1, pp. 48–57, 1996.

[14] 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*, vol. 1807. New York, NY, USA: Springer-Verlag, May 2000, pp. 351–368.

[15] F. Oggier and B. Hassibi, "The secrecy capacity of the MIMO wiretap channel," *IEEE Trans. Inf. Theory*, vol. 57, no. 8, pp. 4961–4972, Aug. 2011.

[16] 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.

[17] A. Khisti and G. W. Wornell, "Secure transmission with multiple antennas—Part II: The MIMOME wiretap channel," *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5515–5532, Nov. 2010.

[18] T. Liu and S. Shamai (Shitz), "A note on the secrecy capacity of the multiple-antenna wiretap channel," *IEEE Trans. Inf. Theory*, vol. 55, no. 6, pp. 2547–2553, Jun. 2009.

[19] 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.

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

[21] Y. Liang, G. Kramer, H. V. Poor, and S. Shamai (Shitz), "Compound wiretap channels," *EURASIP J. Wireless Commun. Netw.*, 2009, Article ID. 142374.

[22] I. Bjelaković, H. Boche, and J. Sommerfeld, "Secrecy results for compound wiretap channels," *Probl. Inf. Transmission*, vol. 49, no. 1, pp. 73–98, Mar. 2013.

[23] 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.

[24] A. Khisti, "Interference alignment for the multiantenna compound wiretap channel," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 2976–2993, May 2011.

[25] R. F. Schaefer and S. Loyka, "The secrecy capacity of a compound MIMO Gaussian channel," in *Proc. IEEE Inf. Theory Workshop*, Seville, Spain, Sep. 2013, pp. 104–108.

[26] R. F. Schaefer and S. Loyka, "The compound secrecy capacity of a class of non-degraded MIMO Gaussian channels," in *Proc. 52th Annual Allerton Conf. Commun., Control, Computing*, Monticello, IL, USA, Oct. 2014, pp. 1004–1010.

[27] R. F. Schaefer and S. Loyka, "On the secrecy capacity of rank-deficient compound wiretap channels," in *Proc. IEEE Global Commun. Conf.—Workshop Trusted Commun. Physical Layer Security*, San Diego, CA, USA, Dec. 2015.

[28] R. F. Schaefer and S. Loyka, "The secrecy capacity of compound MIMO Gaussian channels," *IEEE Trans. Inf. Theory*, to be published, DOI: 10.1109/TIT.2015.2458856.

[29] R. F. Schaefer and H. V. Poor, "Robust transmission over wiretap channels with secret keys," in *Proc. Asilomar Conf. Signals, Systems, Computers*, Pacific Grove, CA, USA, Nov. 2014, pp. 60–64.

[30] D. Blackwell, L. Breiman, and A. J. Thomasian, "The capacities of certain channel classes under random coding," *Ann. Math. Stat.*, vol. 31, no. 3, pp. 558–567, 1960.

[31] R. Ahlswede, "Elimination of correlation in random codes for arbitrarily varying channels," *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, vol. 44, pp. 159–175, 1978.

[32] I. Csiszár and P. Narayan, "The capacity of the arbitrarily varying channel revisited: Positivity, constraints," *IEEE Trans. Inf. Theory*, vol. 34, no. 2, pp. 181–193, Mar. 1988.

[33] E. MolavianJazi, M. Bloch, and J. N. Laneman, "Arbitrary jamming can preclude secure communication," in *Proc. 47th Annual Allerton Conf. Commun., Control, Computing*, Monticello, IL, USA, Sep. 2009, pp. 1069–1075.

[34] I. Bjelaković, H. Boche, and J. Sommerfeld, "Capacity Results for Arbitrarily Varying Wiretap Channels," in *Information Theory, Combinatorics, and Search Theory*. New York, NY, USA: Springer-Verlag, 2013, pp. 123–144.

[35] H. Boche and R. F. Schaefer, "Capacity results and super-activation for wiretap channels with active wiretappers," *IEEE Trans. Inf. Forensics Security*, vol. 8, no. 9, pp. 1482–1496, Sep. 2013.

[36] H. Boche, R. F. Schaefer, and H. V. Poor, "On arbitrarily varying wiretap channels for different classes of secrecy measures," in *Proc. IEEE Int. Symp. Inf. Theory*, Honolulu, HI, USA, Jun. 2014, pp. 2376–2380.

[37] M. Wiese, J. Nötzel, and H. Boche, "The arbitrarily varying wiretap channel—Deterministic and correlated random coding capacities under the strong secrecy criterion," in *Proc. IEEE Int. Symp. Inf. Theory*, Hong Kong, Jun. 2015. [Online]. Available: http://arxiv.org/abs/1410.8078

[38] J. Nötzel, M. Wiese, and H. Boche, "The arbitrarily varying wiretap channel—Secret randomness, stability and super-activation," in *Proc. IEEE Int. Symp. Inf. Theory*, Hong Kong, Jun. 2015. [Online]. Available: http://arxiv.org/abs/1501.07439

[39] H. Boche, R. F. Schaefer, and H. V. Poor, "On the continuity of the secrecy capacity of compound and arbitrarily varying wiretap channels," *IEEE Trans. Inf. Forensics Security*, to be published, DOI: 10.1109/TIFS.2015.2465937.

[40] R. F. Schaefer, H. Boche, and H. V. Poor, "Super-activation as a unique feature of secure communication in malicious environments, Jul. 2015, under review.

[41] G. Smith, J. A. Smolin, and J. Yard, "Quantum communication with Gaussian channels of zero quantum capacity," *Nature Photonics*, vol. 5, no. 10, pp. 624–627, Oct. 2011.

[42] G. Giedke and M. M. Wolf, "Quantum communication: Super-activated channels," *Nature Photonics*, vol. 5, no. 10, pp. 578–580, Oct. 2011.

[43] B. A. Bash, D. Goeckel, and D. Towsley, "Limits of reliable communication with low probability of detection on AWGN channels," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 9, pp. 1921–1930, Sep. 2013.

[44] P. H. Che *et al.*, "Reliable, deniable and hidable communication: A quick survey," in *Proc. IEEE Inf. Theory Workshop*, Hobart, Australia, Nov. 2014, pp. 227–231.

[45] M. R. Bloch, "Covert communication over noisy channels: A resolvability perspective," *IEEE Trans. Inf. Theory*, 2015. [Online]. Available: http://arxiv.org/abs/1503.08778

[46] H. Boche, R. F. Schaefer, A. Grigorescu, and H. V. Poor, Robust code design for wiretap and broadcast channels under channel uncertainty, in preparation.

[47] M. Kobayashi, Y. Liang, S. Shamai (Shitz), and M. Debbah, "On the compound MIMO broadcast channels with confidential messages," in *Proc. IEEE Int. Symp. Inf. Theory*, Seoul, Korea, Jun. 2009, pp. 1283–1287.

[48] R. F. Schaefer and H. Boche, "Robust broadcasting of common and confidential messages over compound channels: Strong secrecy and decoding performance," *IEEE Trans. Inf. Forensics Security*, vol. 9, no. 10, pp. 1720–1732, Oct. 2014.

[49] A. Grigorescu, H. Boche, R. F. Schaefer, and H. V. Poor, "Capacity region continuity of the compound broadcast channel with confidential messages," in *Proc. IEEE Inf. Theory Workshop*, Jerusalem, Israel, Apr. 2015, pp. 1–5.

[50] R. F. Schaefer and H. V. Poor, "On secure communication over multiple access wiretap channels under channel uncertainty," in *Proc. IEEE Conf. Commun. Netw. Security Workshops*, San Francisco, CA, USA, Oct. 2014, pp. 109–114.

[51] H. Zivari-Fard, B. Akhbari, M. Ahmadian-Attari, and M. R. Aref, "Compound multiple access channel with confidential messages," in *Proc. IEEE Int. Conf. Commun.*, Sydney, Australia, Jun. 2014, pp. 1922–1927.

[52] I. Devetak, "The private classical capacity and quantum capacity of a quantum channel," *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 44–55, Jan. 2005.

[53] M. R. Bloch and J. N. Laneman, "Strong secrecy from channel resolvability," *IEEE Trans. Inf. Theory*, vol. 59, no. 12, pp. 8077–8098, Dec. 2013.

[54] J. Hou and G. Kramer, "Effective secrecy: Reliability, confusion, and stealth," in *Proc. IEEE Int. Symp. Inf. Theory*, Honolulu, HI, USA, Jun. 2014, pp. 601–605.

[55] J. L. Massey, "A simplified treatment of Wyners wire-tap channel," in *Proc. 21st Allerton Conf. on Comm., Control and Computing*, Monticello, IL, USA, Oct. 1983, pp. 268–276.

[56] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge, U.K.: Cambridge Univ. Press, 2011.

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

[58] J. Wolfowitz, *Coding Theorems of Information Theory* 3rd ed., New York, NY, USA: Springer-Verlag, 1978.

[59] R. Ahlswede, "A note on the existence of the weak capacity for channels with arbitrarily varying channel probability functions and its relation to Shannon's zero error capacity," *Ann. Math. Stat.*, vol. 41, no. 3, pp. 1027–1033, 1970.

[60] H. Yamamoto, "Rate-distortion theory for the Shannon cipher system," *IEEE Trans. Inf. Theory*, vol. 43, no. 3, pp. 827–835, May 1997.

[61] N. Merhav, "Shannon's secrecy system with informed receivers and its application to systematic coding for wiretapped channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2723–2734, Jun. 2008.

[62] W. Kang and N. Liu, "Wiretap channel with shared key," in *Proc. IEEE Inf. Theory*

Workshop, Dublin, Ireland, Aug. 2010, pp. 1–5.

[63] R. F. Schaefer and A. Khisti, "Secure broadcasting of a common message with independent secret keys," in *Proc. Conf. Inf. Sciences and Systems*, Princeton, NJ, USA, Mar. 2014, pp. 1–6.

[64] R. F. Schaefer, A. Khisti, and H. V. Poor, "How to use independent secret keys for secure broadcasting of common messages," in *Proc. IEEE Int. Symp. Inf. Theory*, Hong Kong, China, Jun. 2015, pp. 1971–1975.

[65] C. E. Shannon, "The zero error capacity of a noisy channel," *IRE Trans. Inf. Theory*, vol. 2, no. 3, pp. 8–19, Sep. 1956.

[66] N. Alon, "The Shannon capacity of a union," *Combinatorica*, vol. 18, no. 3, pp. 301–310, Mar. 1998.

ABOUT THE AUTHORS

Rafael F. Schaefer (Member, IEEE) received the Dipl.-Ing. degree in electrical engineering and computer science from the Technische Universität Berlin, Berlin, Germany, in 2007, and the Dr.-Ing. degree in electrical engineering from the Technische Universität München, Munich, Germany, in 2012.

He was a Research and Teaching Assistant with the Heinrich-Hertz-Lehrstuhl für Mobilkommunikation, Technische Universität Berlin, from 2007 to 2010, and the Lehrstuhl für Theoretische Informationstechnik, Technische Universität München, from 2010 to 2013. He is currently a Postdoctoral Research Fellow with the Department of Electrical Engineering, Princeton University, Princeton, NJ, USA.

Mr. Schaefer was a recipient of the VDE Johann-Philipp-Reis Prize in 2013. He was one of the exemplary reviewers of the IEEE Communication Letters in 2013. Currently, he is an Associate Member of the IEEE Information Forensics and Security Technical Committee.

Holger Boche (Fellow, IEEE) received the Dipl.-Ing. and Dr.-Ing. degrees in electrical engineering from the Technische Universität Dresden, Dresden, Germany, in 1990 and 1994, respectively. He graduated in mathematics from the Technische Universität Dresden in 1992. From 1994 to 1997, he did Postgraduate studies in mathematics at the Friedrich-Schiller Universität Jena, Jena, Germany, and received the Ph.D. degree in pure mathematics from the Technische Universität Berlin, Berlin, Germany, in 1998.

In 1997, he joined the Heinrich-Hertz-Institut (HHI) für Nachrichtentechnik Berlin, Berlin, Germany. Starting in 2002, he was a Full Professor for mobile communication networks with the Institute for Communications Systems, Technische Universität Berlin. In 2003, he became Director of the Fraunhofer German-Sino Lab for Mobile Communications, Berlin, Germany, and in 2004, he became the Director of the Fraunhofer Institute for Telecommunications (HHI), Berlin, Germany. Since October 2010, he has been with the Institute of Theoretical Information Technology and Full Professor at the Technische Universität München, Munich, Germany. Since 2014, he has been a Member and Honorary Fellow of the TUM Institute for Advanced Study, Munich, Germany. He was a Visiting Professor with the ETH Zurich, Zurich, Switzerland, during the 2004 and 2006 Winter terms, and with KTH Stockholm, Stockholm, Sweden, during the 2005 Summer term.

Prof. Boche is a Member of IEEE Signal Processing Society SPCOM and SPTM Technical Committee. He was elected a Member of the German Academy of Sciences (Leopoldina) in 2008 and of the Berlin Brandenburg Academy of Sciences and Humanities in 2009. He received the Research Award "Technische Kommunikation" from the Alcatel SEL Foundation in October 2003, the "Innovation Award" from the Vodafone Foundation in June 2006, and the Gottfried Wilhelm Leibniz Prize from the Deutsche Forschungsgemeinschaft (German Research Foundation) in 2008. He was corecipient of the 2006 IEEE Signal Processing Society Best Paper Award and recipient of the 2007 IEEE Signal Processing Society Best Paper Award.

H. Vincent Poor (Fellow, IEEE) received the Ph.D. degree in electrical engineering and computer science from Princeton University, Princeton, NJ, USA, in 1977.

From 1977 until 1990, he was on the Faculty of the University of Illinois at Urbana-Champaign. Since 1990, he has been on the faculty at Princeton, where he is the Dean of Engineering and Applied Science, and the Michael Henry Strater University Professor of Electrical Engineering. He has also held visiting appointments at several other institutions, most recently at Imperial College and Stanford. Dr. Poor's research interests are in the areas of information theory, stochastic analysis and statistical signal processing, and their applications in wireless networks and related fields. Among his publications in these areas is the recent book *Mechanisms and Games for Dynamic Spectrum Allocation* (Cambridge University Press, 2014).

Dr. Poor is a member of the National Academy of Engineering and the National Academy of Sciences, and is a foreign member of Academia Europaea and the Royal Society. He is also a fellow of the American Academy of Arts and Sciences, the Royal Academy of Engineering (U.K.), and the Royal Society of Edinburgh. In 1990, he served as President of the IEEE Information Theory Society, and in 2004–07 as the Editor-in-Chief of the IEEE Transactions on Information Theory. He received a Guggenheim Fellowship in 2002 and the IEEE Education Medal in 2005. Recent recognition of his work includes the 2014 URSI Booker Gold Medal, and honorary doctorates from Aalborg University, Aalto University, HKUST and the University of Edinburgh.