Robust Broadcasting of Common and Confidential Messages Over Compound Channels: Strong Secrecy and Decoding Performance

Rafael F. Schaefer, *Member, IEEE*, and Holger Boche, *Fellow, IEEE*

***Abstract*— The broadcast channel with confidential messages (BCC) consists of one transmitter and two receivers, where the transmitter sends a common message to both receivers and, at the same time, a confidential message to one receiver which has to be kept secret from the other one. In this paper, this communication scenario is studied for compound channels, where it is only known to the transmitter and receivers that the actual channel realization is fixed and from a prespecified set of channels. The information theoretic criterion of strong secrecy is analyzed in detail and its impact on the decoding performance of the non-legitimate receiver is characterized. In particular, it is shown that regardless of the computational capabilities and the applied decoding strategy of the non-legitimate receiver, his decoding error always tends to one. This gives a valuable signal processing implication of the strong secrecy criterion and identifies desirable properties of an optimal code design. Further, an achievable strong secrecy rate region is derived and a multiletter outer bound is given. Both together yield a multiletter expression of the strong secrecy capacity region of the compound BCC.**

***Index Terms*— Broadcast channel with confidential messages, secrecy capacity, strong secrecy, compound channel, decoding performance, embedded security.**

I. Introduction

RAPID developments in communication systems make information available almost everywhere. Along with this, the security of sensitive information from unauthorized access becomes an important task. Especially wireless communication systems are inherently vulnerable, since transmitted signals are received by intended users but are also easily eavesdropped by non-legitimate receivers.

Nowadays, the architecture of communication systems is usually separated into error correction and data encryption. The error correction is typically implemented on the physical layer turning the noisy communication channel into an ideal "*bit pipe*." Then on higher layers, the data encryption is realized by applying cryptographic techniques which are based on the assumption of insufficient computational capabilities of non-legitimate receivers. Accordingly, one is interested in unconditional security which does not rely on such limitations.

Another approach for realizing security is the so-called concept of information theoretic secrecy. This was initiated by Shannon in his seminal paper [3]. He showed that a secret key used as *one-time pad* allows for secure communication over a noiseless channel. Subsequently, Wyner introduced the *wiretap channel* [4], which describes the communication scenario over a noisy channel and without any secret keys. The key insight of this work is that information theoretic secrecy can be achieved by exploiting the physical properties of the noisy channel. This was later generalized by Csiszár and Körner to the *broadcast channel with confidential messages* [5]. In these works, the secrecy of confidential information was measured using the criterion of weak secrecy. This area of research has drawn attention in recent years as it provides a promising approach to embed secure communication in wireless networks; for instance see [6]–[9] and references therein. Concurrently, it has been demonstrated that the joint implementation of different public and secure communication tasks on the physical layer can lead to significant gains in spectral efficiency. Thus, it is not surprising that the efficient *physical layer service integration* [10] has also been identified by operators and national agencies as a promising task to increase the efficiency of public communication systems [11], [12].

However, as already mentioned, the criterion of *weak secrecy* is usually applied, which is heuristic in nature in that no convincing operational meaning has been given to it yet. This means that even if this criterion holds, it is not clear how secure the confidential information actually is. But recently, an operational meaning has been given to the *strong secrecy* criterion introduced by Csiszár [13] and by Maurer and Wolf [14]: it was established in [15] for the wiretap channel that the strong secrecy criterion implies that the average decoding error at the eavesdropper tends to one exponentially fast for any decoder he (or she) may use. This gives the strong secrecy criterion important signal processing consequences and therewith paves the way for providing secure communication with guaranteed, i.e., provable, secrecy.

Another challenge, especially for operators of wireless communication systems, is the provision of accurate channel state information at transmitter and receivers. Practical systems always suffer from channel uncertainty due to the nature

Manuscript received August 20, 2013; revised June 19, 2014; accepted August 9, 2014. Date of publication August 14, 2014; date of current version September 12, 2014. The work of R. F. Schaefer was supported by the German Research Foundation under Grant WY-151/2-1. The work of H. Boche was supported in part by the German Ministry of Education and Research under Grant 01BQ1050 and in part by the German Research Foundation under Grant BO-1734/25-1. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Giuseppe Persiano.

R. F. Schaefer is with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544 USA (e-mail: rafaelfs@princeton.edu).

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

Digital Object Identifier 10.1109/TIFS.2014.2348193

1556-6013 © 2014 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

of the wireless medium and estimation/feedback inaccuracy. In addition, it is hard to believe that non-legitimate receivers will share their channel information with the transmitter making eavesdropping even harder. Thus, it is reasonable to assume that the exact channel realization is not known; rather, it is only known that it belongs to a pre-specified set of channels. If this channel remains fixed during the whole transmission of a codeword, this corresponds to the concept of *compound channels* [16], [17]. Accordingly, one is interested in robust strategies that allow for secure communication over compound channels. The compound wiretap channel is analyzed for discrete memoryless channels in [15] and [18] and for MIMO Gaussian channels in [19] and [20]. First studies for the MIMO Gaussian compound broadcast channel with confidential messages can be found in [21].

In this paper we consider the discrete memoryless *compound broadcast channel with confidential messages (BCC)*. This models the communication scenario with one transmitter and two receivers, where the transmitter broadcasts a common message to both receivers and, at the same time, sends a confidential message to receiver 1 which has to be kept secret from receiver 2. Thus, receiver 2 is both a legitimate receiver for the common message and a non-legitimate receiver for the confidential message. This is in contrast to the classical wiretap channel [4], where the eavesdropper is solely a non-legitimate receiver and does not belong to the communication system. The corresponding system model for the compound BCC is introduced in detail in Section II.

We start with the classical approach and measure the secrecy of the confidential message using the information theoretic concept of strong secrecy. As the non-legitimate receiver is part of the communication system, he may have further side information available to infer the confidential message; most obvious the common message he is intended to decode. Hence, it is reasonable to question the validity of the classical strong secrecy criterion. Accordingly, in Section III we generalize the information theoretic secrecy criterion by taking such side information into account. Furthermore, we also analyze the secrecy of the confidential message from a signal processing point of view by characterizing the decoding performance of the non-legitimate receiver. We show that the information theoretic criterion of strong secrecy and the signal processing inspired criterion of worst decoding performance are connected which gives the strong secrecy criterion an important signal processing interpretation. In addition, we identify *vanishing output variation* at the non-legitimate receiver as a desirable code property since it turns out that such codes realize secrecy for all the discussed information theoretic and signal processing criteria simultaneously.

The confidential message must be protected against the non-legitimate receiver 2 requiring a code which reveals no information to him. But simultaneously, he is also a legitimate receiver since the common message must be successfully transmitted to him requiring a code suitable for reliable communication. At a first glance, these two goals seem to be conflictive making the code design for the compound BCC a challenging task. In Section IV an achievable secrecy rate region is established. Thereby it is shown that codes having the vanishing output variation property are further suitable to simultaneously meet both conflicting intentions of secrecy and reliability. A multi-letter outer bound is then given in Section V, which establishes, together with the achievable rate region, a multi-letter characterization of the secrecy capacity region. Finally, a conclusion is given in Section VI.

Notation

Discrete random variables are denoted by capital letters and their realizations and ranges by lower case and script letters, respectively; $X - Y - Z$ denotes a Markov chain of random variables X, Y, and Z in this order; all logarithms and information quantities are taken to the base 2; $\mathbb{N}$ and $\mathbb{R}_+$ are the sets of non-negative integers and non-negative real numbers; $\mathcal{A}^c$, $|\mathcal{A}|$, and $\mathcal{A}\times\mathcal{B}$ are the complement, cardinality, and Cartesian product of the sets $\mathcal{A}$ and $\mathcal{B}$; $H(\cdot)$, $I(\cdot;\cdot)$, and $D(\cdot\|\cdot)$ are the traditional entropy, mutual information, and Kullback-Leibler (information) divergence; $\|\mu - \nu\|$ is the total variation distance of measures μ and ν on $\mathcal{A}$ defined as $\|\mu - \nu\| := \sum_{a\in\mathcal{A}} |\mu(a) - \nu(a)|$ or equivalently as $\|\mu - \nu\| := 2\sup_{A\subseteq\mathcal{A}} |\mu(A) - \nu(A)|$, cf. for example [22, Lemma 4.1.1]; $\mathcal{P}(\cdot)$ denotes the set of all probability distributions; the product distribution $P_A P_B$ is defined by the product marginal distributions of its components P_A and P_B, i.e., $P_A P_B(a, b) = P_A(a)P_B(b)$ for all $a \in \mathcal{A}$, $b \in \mathcal{B}$; $\mathbb{E}[\cdot]$ and $\mathbb{P}\{\cdot\}$ are the expectation and probability; $\mathbb{1}_{\mathcal{A}}(\cdot)$ denotes the indicator function, i.e., $\mathbb{1}_{\mathcal{A}}(a) = 1$ if $a \in \mathcal{A}$ and $\mathbb{1}_{\mathcal{A}}(a) = 0$ otherwise; lhs := rhs means the value of the right hand side (rhs) is assigned to the left hand side (lhs), lhs =: rhs is defined accordingly.

II. Compound Broadcast Channel With Confidential Messages

Here we introduce the system model for the compound BCC. Thereby we start with the classical criterion of information theoretic strong secrecy to measure the secrecy of the confidential message. The communication problem at hand is depicted in Fig. 1 and formalized as follows.

Let $\mathcal{X}$ and $\mathcal{Y}$, $\mathcal{Z}$ be finite input and output sets and $\mathcal{S}$ be a finite state set. Then for fixed channel realization $s \in \mathcal{S}$ and input and output sequences $x^n \in \mathcal{X}^n$ and $y^n \in \mathcal{Y}^n$, $z^n \in \mathcal{Z}^n$ of length n, the discrete memoryless broadcast channel is given by $W_s^n(y^n, z^n|x^n) := \prod_{i=1}^n W_s(y_i, z_i|x_i)$. We denote its marginal channels by $W_{\mathcal{Y},s}^n(y^n|x^n)$ and $W_{\mathcal{Z},s}^n(z^n|x^n)$.

Definition 1: The discrete memoryless *compound broadcast channel* $\mathfrak{W}$ is given by the families of pairs of channels with common input as

$$\mathfrak{W} := \big\{(W_{\mathcal{Y},s}, W_{\mathcal{Z},s}) : s \in \mathcal{S}\big\}.$$

Remark 1: This includes the widely adopted model of the form $\mathfrak{W} = \{(W_{\mathcal{Y},s}, W_{\mathcal{Z},t}) : s \in \mathcal{S}, t \in \mathcal{T}\}$ with $\mathcal{S} \neq \mathcal{T}$ as we can always construct a new set of the form $\hat{S} = \mathcal{S} \times \mathcal{T}$.

We consider a block code of arbitrary but fixed length n. Let $\mathcal{M}_0 := \{1, ..., M_{0,n}\}$ and $\mathcal{M}_1 := \{1, ..., M_{1,n}\}$ be the sets of common and confidential messages. In addition, we use the abbreviation $\mathcal{M} := \mathcal{M}_0 \times \mathcal{M}_1$.

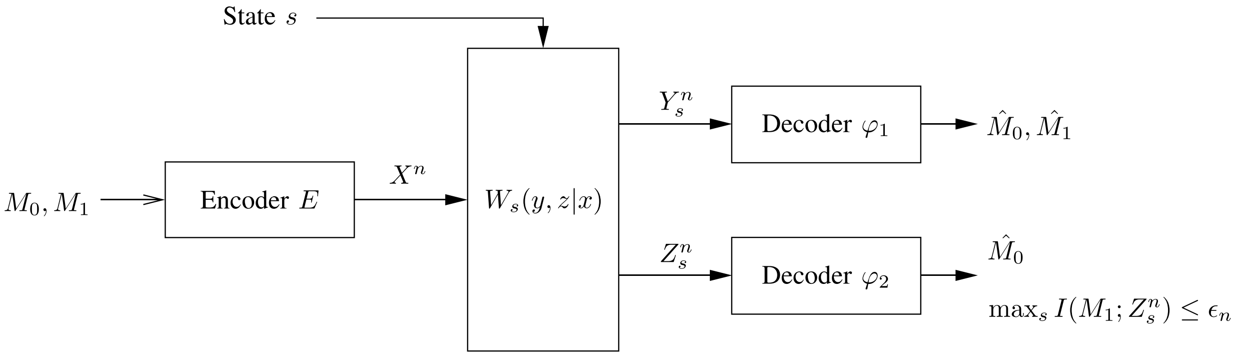

Fig. 1. Compound broadcast channel with confidential messages (BCC). The transmitter encodes the message pair (M_0, M_1) into the codeword $X^n = E(M_0, M_1)$ and transmits it over the compound broadcast channel $\mathfrak{W}$ to the receivers, which have to decode their intended messages $(\hat{M}_0, \hat{M}_1) = \varphi_1(Y_s^n)$ and $\hat{M}_0 = \varphi_2(Z_s^n)$ for any channel realization $s \in \mathcal{S}$. At the same time, receiver 2 has to be kept ignorant of M_1 in the sense that $\max_{s\in\mathcal{S}} I(M_1; Z_s^n) \le \epsilon_n$.

Definition 2: An $(n, M_{0,n}, M_{1,n})$*-code* for the compound BCC consists of a stochastic encoder

$$E : \mathcal{M}_0 \times \mathcal{M}_1 \to \mathcal{P}(\mathcal{X}^n) \tag{1}$$

i.e., a stochastic matrix, and decoders at receivers 1 and 2

$$\varphi_1 : \mathcal{Y}^n \to \mathcal{M}_0 \times \mathcal{M}_1 \tag{2a}$$

$$\varphi_2 : \mathcal{Z}^n \to \mathcal{M}_0. \tag{2b}$$

The encoder is allowed to be stochastic in the sense that it is specified by conditional probabilities $E(x^n|m_0, m_1)$ with $\sum_{x^n\in\mathcal{X}^n} E(x^n|m_0, m_1) = 1$ for each $m_0 \in \mathcal{M}_0$ and $m_1 \in \mathcal{M}_1$. Thus, $E(x^n|m_0, m_1)$ denotes the probability that the message pair $(m_0, m_1) \in \mathcal{M}_0 \times \mathcal{M}_1$ is encoded as the codeword $x^n \in \mathcal{X}^n$.

Remark 2: Already for the classical wiretap channel (without additional common message) it is shown that a stochastic encoder is needed to guarantee the secrecy of the confidential message, see [9, Sec. 3.4]. Although the definition of a stochastic encoder is given in a very general form, it turns out that a much easier stochastic structure will be sufficient. For details we refer to Section IV, where the specific structure of the stochastic encoder is presented.

Remark 3: At some points it is beneficial to express the decoders in terms of decoding sets. Then φ_1 in (2a) is specified by disjoint decoding sets $\{\mathcal{D}_1(m_0, m_1) \subset \mathcal{Y}^n : (m_0, m_1) \in \mathcal{M}_0 \times \mathcal{M}_1\}$ and, similarly, φ_2 in (2b) by disjoint decoding sets $\{\mathcal{D}_2(m_0) \subset \mathcal{Z}^n : m_0 \in \mathcal{M}_0\}$.

Encoder and decoders have to be designed in such a way that they realize reliable communication and secrecy simultaneously. Moreover, since neither the transmitter nor the receivers know the actual channel realization, they must be universal such that they work for all channel realizations simultaneously. The communication task of reliable transmission of all messages to their respective receivers over the compound broadcast channel is addressed first.

When the transmitter has sent the message pair $m = (m_0, m_1) \in \mathcal{M}$ and receivers 1 and 2 have received $y^n \in \mathcal{Y}^n$ and $z^n \in \mathcal{Z}^n$, the decoder at receiver 1 is in error if $y^n \notin \mathcal{D}_1(m_0, m_1)$. Accordingly, the decoder at receiver 2 is in error if $z^n \notin \mathcal{D}_2(m_0)$. Then for an $(n, M_{0,n}, M_{1,n})$-code, the average probabilities of error at receivers 1 and 2 for channel realization $s \in \mathcal{S}$ are then given by

$$\bar{e}_{1,n}(s) := \frac{1}{|\mathcal{M}|}\sum_{m\in\mathcal{M}}\sum_{x^n\in\mathcal{X}^n} W^n_{\mathcal{Y},s}(\mathcal{D}_1^c(m_0, m_1)|x^n)E(x^n|m)$$

$$\bar{e}_{2,n}(s) := \frac{1}{|\mathcal{M}|}\sum_{m\in\mathcal{M}}\sum_{x^n\in\mathcal{X}^n} W^n_{\mathcal{Z},s}(\mathcal{D}_2^c(m_0)|x^n)E(x^n|m).$$

As reliable communication has to be guaranteed for all $s \in \mathcal{S}$, we set $\bar{e}_{i,n} = \max_{s\in\mathcal{S}} \bar{e}_{i,n}(s)$, $i = 1, 2$.

To ensure the confidential message to be kept secret from non-legitimate receiver 2 for all channel realizations $s \in \mathcal{S}$, we require $\max_{s\in\mathcal{S}} I(M_1; Z_s^n) \le \epsilon_n$ for some (small) $\epsilon_n > 0$ with M_1 the random variable uniformly distributed over the set $\mathcal{M}_1$ and $Z_s^n = (Z_{s,1}, Z_{s,2}, ..., Z_{s,n})$ the output at receiver 2 for channel realization $s \in \mathcal{S}$. This criterion is known as *strong secrecy* [13], [14] and yields the following definition.

Definition 3: A rate pair $(R_0, R_1) \in \mathbb{R}_+^2$ is said to be *achievable* for the compound BCC $\mathfrak{W}$ if for any $\tau > 0$ there is an $n(\tau) \in \mathbb{N}$ and a sequence of $(n, M_{0,n}, M_{1,n})$-codes such that for all $n \ge n(\tau)$ we have $\frac{1}{n}\log M_{0,n} \ge R_0 - \tau$, $\frac{1}{n}\log M_{1,n} \ge R_1 - \tau$, and

$$\max_{s\in\mathcal{S}} I(M_1; Z_s^n) \le \epsilon_n := 2^{-n\alpha} \tag{3}$$

for some $\alpha > 0$ while $\bar{e}_{1,n}, \bar{e}_{2,n}, \epsilon_n \to 0$ as $n \to \infty$. The set of all achievable rate pairs is the *strong secrecy capacity region* of the compound BCC $\mathfrak{W}$ and is denoted by $\mathcal{C}_S(\mathfrak{W})$.

Remark 4: Note that the explicit requirement of an exponentially fast decreasing ϵ_n of the form $2^{-n\alpha}$, $\alpha > 0$, in (3) is no restriction. There will be no loss of generality since all applied methods will actually provide an exponentially fast decrease.

The secrecy of the confidential message is characterized by the mutual information between the confidential message and the channel output at the non-legitimate receiver. This is the classical approach for measuring secrecy and was already used in [5] for the classical BCC. However, there were no implications on the non-legitimate receiver discussed. In particular, it is not clear what he can or cannot do to infer the confidential information, especially since he is part of the communication system. Moreover, as he is supposed to decode the common message, this message might provide some knowledge about the confidential message. Accordingly, the secrecy criterion should reflect this fact. These questions will be addressed in the following section and it will be shown that the secrecy criterion should be generalized to include the knowledge about the common message as well.

III. Strong Secrecy and Decoding Performance

Here we want to analyze the strong secrecy criterion and the desirable decoding performance of the non-legitimate receiver in more detail.

A. Strong Secrecy and Vanishing Output Variation

The concept of *vanishing output variation* has been identified to be necessary for achieving the strong secrecy capacity of the wiretap channel with side information at the eavesdropper [23]. It suggests itself to study this concept also for the compound BCC. For this purpose we define for each $s \in \mathcal{S}$ the channel to the non-legitimate receiver

$$\overline{W}^n_{\mathcal{Z},s}(z^n|m_0,m_1) := \sum_{x^n\in\mathcal{X}^n} W^n_{\mathcal{Z},s}(z^n|x^n)E(x^n|m_0,m_1) \quad (4)$$

which takes the stochastic encoder E into account, cf. (1).

Definition 4: A code for the compound BCC has exponentially fast *vanishing output variation* if there exists for each channel realization $s \in \mathcal{S}$ and each common message $m_0 \in \mathcal{M}_0$ a measure[1] ϑ_{s,m_0} on $\mathcal{Z}^n$ such that for all $m_1 \in \mathcal{M}_1$ it holds

$$\sum_{z^n\in\mathcal{Z}^n} \left|\overline{W}^n_{\mathcal{Z},s}(z^n|m_0,m_1) - \vartheta_{s,m_0}(z^n)\right| \leq 2^{-n\beta} \quad (5)$$

for some $\beta > 0$. Instead of (5) we sometimes write $\|\overline{W}^n_{\mathcal{Z},s}(\cdot|m_0,m_1) - \vartheta_{s,m_0}\| \leq 2^{-n\beta}$ interchangeably.

Intuitively, Definition 4 has the following meaning: From a secrecy perspective, the safest option is to assume that the non-legitimate receiver 2 is aware the common message $m_0 \in \mathcal{M}_0$ (he is supposed to decode it anyway) and the actual channel realization $s \in \mathcal{S}$ (due to potential channel estimation, etc.). Then the property of vanishing output variation (5) ensures that regardless of which common message $m_0 \in \mathcal{M}_0$ is transmitted and which channel realization $s \in \mathcal{S}$ controls the channel, the output at the receiver 2 "looks" the same. In more detail, for all potentially transmitted messages $m_1 \in \mathcal{M}_1$, the channel output at receiver 2 is, basically, given by ϑ_{s,m_0}, which is independent of $m_1 \in \mathcal{M}_1$. Thus, the non-legitimate receiver will not be able to learn anything meaningful about the confidential message from his channel output.

Now, the following result shows that vanishing output variation (5) implies strong secrecy (3), which establishes an desirable and important property of such codes.

Proposition 1: If a code for the compound BCC has the vanishing output variation property, then the strong secrecy criterion satisfies

$$\max_{s\in\mathcal{S}} I(M_1;Z^n_s) \leq \epsilon_n \quad (6)$$

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

Proof: Let $P_{Z^n_s M_0 M_1}$ be the joint distribution and P_{M_0}, P_{M_1}, and $P_{Z^n_s}$ be the corresponding marginal distributions where the former are uniformly distributed over the sets of messages $\mathcal{M}_0$ and $\mathcal{M}_1$, respectively. With this and $P_{Z^n_s M_1}(z^n,m_1) = \sum_{m_0\in\mathcal{M}_0} P_{Z^n_s M_0 M_1}(z^n,m_0,m_1)$ we observe that we can write

$$P_{Z^n_s M_1}(z^n,m_1) = \frac{1}{|\mathcal{M}|}\sum_{m_0\in\mathcal{M}_0} \overline{W}^n_{\mathcal{Z},s}(z^n|m_0,m_1). \quad (7)$$

If the code has the vanishing output variation property, we then have for each $s \in \mathcal{S}$ and $m_0 \in \mathcal{M}_0$ a measure ϑ_{s,m_0} which satisfies (5). Averaging over all common messages we obtain the measure

$$\bar{\vartheta}_s(z^n) := \frac{1}{|\mathcal{M}_0|}\sum_{m_0\in\mathcal{M}_0} \vartheta_{s,m_0}(z^n)$$

so that

$$\begin{aligned}\bar{\vartheta}_{s,M_1}(z^n,m_1) &:= \frac{1}{|\mathcal{M}_1|}\bar{\vartheta}_s(z^n)\\ &= \frac{1}{|\mathcal{M}|}\sum_{m_0\in\mathcal{M}_0} \vartheta_{s,m_0}(z^n) \quad (8)\end{aligned}$$

defines a product measure on $\mathcal{Z}^n \times \mathcal{M}_1$. Now by the triangle inequality we can bound the total variation distance by[2]

$$\begin{aligned}&\|P_{Z^n_s M_1} - P_{Z^n_s}P_{M_1}\|\\ &\quad\leq \|P_{Z^n_s M_1} - \bar{\vartheta}_{s,M_1}\| + \|\bar{\vartheta}_{s,M_1} - P_{Z^n_s}P_{M_1}\|. \quad (9)\end{aligned}$$

Next we bound both terms individually. With (7) and (8) we obtain for the first term

$$\begin{aligned}&\|P_{Z^n_s M_1} - \bar{\vartheta}_{s,M_1}\|\\ &\quad= \sum_{m_1\in\mathcal{M}_1}\sum_{z^n\in\mathcal{Z}^n}\left|P_{Z^n_s M_1}(z^n,m_1) - \bar{\vartheta}_{s,M_1}(z^n,m_1)\right|\\ &\quad= \sum_{m_1\in\mathcal{M}_1}\sum_{z^n\in\mathcal{Z}^n}\Bigg|\frac{1}{|\mathcal{M}|}\sum_{m_0\in\mathcal{M}_0}\overline{W}^n_{\mathcal{Z},s}(z^n|m_0,m_1)\\ &\qquad\qquad - \frac{1}{|\mathcal{M}|}\sum_{m_0\in\mathcal{M}_0}\vartheta_{s,m_0}(z^n)\Bigg|\\ &\quad\leq \frac{1}{|\mathcal{M}|}\sum_{m\in\mathcal{M}}\sum_{z^n\in\mathcal{Z}^n}\left|\overline{W}^n_{\mathcal{Z},s}(z^n|m_0,m_1) - \vartheta_{s,m_0}(z^n)\right|\\ &\quad\leq 2^{-n\beta} \quad (10)\end{aligned}$$

where the last step follows from the vanishing output variation property, cf. (5).

[1] A measure ϑ on $\mathcal{Z}^n$ is assumed to satisfy the standard properties of non-negativity, i.e., $\vartheta(\mathcal{A}) \geq 0$ for all $\mathcal{A} \subseteq \mathcal{Z}^n$, null empty set, i.e., $\vartheta(\emptyset) = 0$, and countable additivity, i.e., for all collections $\{\mathcal{A}_i\}_{i\in\mathcal{I}}$ of pairwise disjoint sets it holds $\vartheta(\bigcup_{i\in\mathcal{I}}\mathcal{A}_i) = \sum_{i\in\mathcal{I}}\vartheta(\mathcal{A}_i)$. We do not require $\vartheta(\mathcal{Z}^n) = 1$, i.e., ϑ is not necessarily a probability measure.

[2] Note that the total variation distance in (9) is defined on $\mathcal{Z}^n \times \mathcal{M}_1$. This is in contrast to the vanishing output variation property in (5), which is defined on $\mathcal{Z}^n$ only.

Similarly, with (7) and (8) we get for the second term

$$\begin{aligned}
&\|\bar{\vartheta}_{s,M_1} - P_{Z_s^n}P_{M_1}\| \\
&= \sum_{m_1\in\mathcal{M}_1}\sum_{z^n\in\mathcal{Z}^n}\left|\bar{\vartheta}_{s,M_1}(z^n,m_1) - P_{Z_s^n}(z^n)P_{M_1}(m_1)\right| \\
&= \sum_{m_1\in\mathcal{M}_1}\sum_{z^n\in\mathcal{Z}^n}\left|\frac{1}{|\mathcal{M}|}\sum_{i\in\mathcal{M}_0}\vartheta_{s,i}(z^n) - \frac{1}{|\mathcal{M}_1|}P_{Z_s^n}(z^n)\right| \\
&= \sum_{m_1\in\mathcal{M}_1}\sum_{z^n\in\mathcal{Z}^n}\left|\frac{1}{|\mathcal{M}|}\frac{1}{|\mathcal{M}_1|}\sum_{(i,j)\in\mathcal{M}}\vartheta_{s,i}(z^n)\right. \\
&\qquad\qquad\left. -\frac{1}{|\mathcal{M}|}\frac{1}{|\mathcal{M}_1|}\sum_{(i,j)\in\mathcal{M}}\overline{W}^n_{\mathcal{Z},s}(z^n|i,j)\right| \\
&\le \frac{1}{|\mathcal{M}|}\frac{1}{|\mathcal{M}_1|}\sum_{(i,j)\in\mathcal{M}}\sum_{m_1\in\mathcal{M}_1}\sum_{z^n\in\mathcal{Z}^n}\left|\vartheta_{s,i}(z^n)\right. \\
&\qquad\qquad\left. - \overline{W}^n_{\mathcal{Z},s}(z^n|i,j)\right| \\
&\le 2^{-n\beta} \qquad (11)
\end{aligned}$$

where the third step follows from the fact that $\vartheta_{s,i}$ does not depend on $j \in \mathcal{M}_1$ and the last step follows from (5). From (10) and (11) follows that the total variation distance in (9) is exponentially small; more precisely we have

$$\|P_{Z_s^nM_1} - P_{Z_s^n}P_{M_1}\| \le 2\cdot 2^{-n\beta} \qquad (12)$$

for all $s \in \mathcal{S}$. Then the continuity of the entropy function, cf. for example [24, Lemma 1.2.7], implies that the corresponding mutual information term is exponentially small as well, i.e.,

$$\begin{aligned}
I(M_1;Z_s^n) &= H(Z_s^n) + H(M_1) - H(Z_s^n,M_1) \\
&= H(P_{Z_s^n}P_{M_1}) - H(P_{Z_s^nM_1}) \\
&\le -2\cdot 2^{-n\beta}\log(2\cdot 2^{-n\beta}) + 2n\cdot 2^{-n\beta}\log(|\mathcal{Z}||\mathcal{M}_1|) \\
&\le 2^{-n\beta/2} =: \epsilon_n
\end{aligned}$$

for all $s \in \mathcal{S}$ where the last inequality holds for n large enough. This shows (6) and completes the proof. ■

This result shows that vanishing output variation (5) guarantees that the strong secrecy criterion (3) is satisfied. However, for the classical wiretap channel this property further allowed to characterize the decoding performance of the eavesdropper [15], [23]. Moreover, a connection between the strong secrecy criterion to the decoding performance of the eavesdropper was established which gives strong secrecy an operational meaning. Having this in mind, we are interested in establishing similar results for the compound BCC, which is addressed next.

B. Decoding Performance of Non-Legitimate Receiver

The classical approach of measuring the secrecy is done via the mutual information between the confidential message and the channel output at the non-legitimate receiver, i.e., $\max_{s\in\mathcal{S}} I(M_1;Z_s^n) \le \epsilon_n$ as in (3). However, it is not clear how this reflects the capabilities of the non-legitimate receiver. The confidential information has to be protected against non-legitimate receivers on which no assumptions or restrictions are imposed. In particular, there are no restrictions on the computational capabilities or post-processing strategies, which is in contrast to the classical cryptographic approach.

In addition, to characterize the secrecy of the confidential message also from a signal processing point of view, we want to analyze the decoding performance of the non-legitimate receiver as well. To obtain guaranteed performance bounds, one has to prepare for the worst. This is a non-legitimate receiver who knows the actual channel realization $s \in \mathcal{S}$, but also the common message $m_0 \in \mathcal{M}_0$. This is valid, since receiver 2 is part of the communication system and the code will be designed such that he will be able to decode the common message. However, such knowledge must not provide any information about the confidential information. In more detail, knowing $s \in \mathcal{S}$ and $m_0 \in \mathcal{M}_0$, receiver 2 is able to create decoding sets

$$\left\{\widetilde{\mathcal{D}}_{s,m_0}(m_1) \subset \mathcal{Z}^n : m_1 \in \mathcal{M}_1\right\} \qquad (13)$$

with $\bigcup_{m_1\in\mathcal{M}_1}\widetilde{\mathcal{D}}_{s,m_0}(m_1) = \mathcal{Z}^n$ and $\widetilde{\mathcal{D}}_{s,m_0}(m_1)\cap\widetilde{\mathcal{D}}_{s,m_0}(\hat{m}_1) = \emptyset$ for $\hat{m}_1 \neq m_1$. Thus, in contrast to the original communication problem, cf. Definition 2 an Remark 3, we allow the decoding sets to depend on the particular channel realization and common message. For channel realization $s \in \mathcal{S}$ this defines the corresponding average decoding error as

$$\bar{e}'_{2,n}(s) := \frac{1}{|\mathcal{M}_0||\mathcal{M}_1|}\sum_{m_0\in\mathcal{M}_0}\sum_{m_1\in\mathcal{M}_1}\overline{W}^n_{\mathcal{Z},s}(\widetilde{\mathcal{D}}^c_{s,m_0}(m_1)|m_0,m_1).$$

With this, we define the best decoding performance of the non-legitimate receiver as $\bar{e}'_{2,n} = \min_{s\in\mathcal{S}}\bar{e}'_{2,n}(s)$.

Having in mind that the non-legitimate receiver is aware of the common message, it is also reasonable to question the validity of the expression $\max_{s\in\mathcal{S}} I(M_1;Z_s^n) \le \epsilon_n$, cf. (3). Therefore we also want to analyze what happens if we replace this by

$$\max_{s\in\mathcal{S}} I(M_1;Z_s^n|M_0) \le \epsilon_n. \qquad (14)$$

Such a secrecy criterion with conditioning on a common part also appears in [25] in the context of rate-distortion-based secrecy for optical communication.

Remark 5: With M_0 and M_1 independent of each other, the secrecy criterion (14) can equivalently be written as

$$\max_{s\in\mathcal{S}} I(M_1;Z_s^n,M_0) \le \epsilon_n.$$

Such kind of formulation is common in genie-aided upper bounds where some additional information is provided to the receiver.

Fortunately, the following result shows that a code with vanishing output variation (5) implies also strong secrecy in the sense of (14) and further yields the worst behavior of decoding performance at the non-legitimate receiver.

Proposition 2: For any given code of Definition 2, assume that the non-legitimate receiver knows the channel realization $s \in \mathcal{S}$ and the common message $m_0 \in \mathcal{M}_0$ and chooses arbitrary decoding sets as in (13). If the code has vanishing output variation according to Definition 4, cf. (5), then the following holds:

i) The strong secrecy criterion satisfies

$$\max_{s\in\mathcal{S}} I(M_1;Z_s^n|M_0) \le \epsilon_n \qquad (15)$$

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

ii) The average probability of decoding error at the non-legitimate receiver satisfies

$$\min_{s\in\mathcal{S}} \bar{e}'_{2,n}(s) \geq 1 - \frac{1}{|\mathcal{M}_1|} - \lambda_n \tag{16}$$

with $1/|\mathcal{M}_1| \to 0$ and $\lambda_n \to 0$ exponentially fast as $n \to \infty$.

Proof: We start with the first assertion. The proof is similar to the one of Proposition 1 and the one given in [25]. For any $s \in \mathcal{S}$ and $m_0 \in \mathcal{M}_0$ we have

$$\begin{aligned}
&\|P_{Z_s^n M_1|M_0=m_0} - P_{Z_s^n|M_0=m_0} P_{M_1|M_0=m_0}\| \\
&= \sum_{m_1\in\mathcal{M}_1}\sum_{z^n\in\mathcal{Z}^n} \big|P_{Z_s^n M_1|M_0}(z^n, m_1|m_0) \\
&\qquad\qquad - P_{Z_s^n|M_0}(z^n|m_0) P_{M_1|M_0}(m_1|m_0)\big| \\
&= \frac{1}{|\mathcal{M}_1|}\sum_{m_1\in\mathcal{M}_1}\sum_{z^n\in\mathcal{Z}^n} \big|\overline{W}^n_{\mathcal{Z},s}(z^n|m_0,m_1) - P_{Z_s^n|M_0}(z^n|m_0)\big| \\
&\leq \frac{1}{|\mathcal{M}_1|}\sum_{m_1\in\mathcal{M}_1}\sum_{z^n\in\mathcal{Z}^n} \Big(\big|\overline{W}^n_{\mathcal{Z},s}(z^n|m_0,m_1) - \vartheta_{s,m_0}(z^n)\big| \\
&\qquad\qquad + \big|\vartheta_{s,m_0}(z^n) - P_{Z_s^n|M_0}(z^n|m_0)\big|\Big)
\end{aligned}$$

where the steps follow from the definition of total variation distance defined on $\mathcal{Z}^n \times \mathcal{M}_1$, the fact that M_0 and M_1 are independent, and from the triangle inequality.

From (5) we know that $\sum_{z^n\in\mathcal{Z}^n} |\overline{W}^n_{\mathcal{Z},s}(z^n|m_0,m_1) - \vartheta_{s,m_0}(z^n)| \leq 2^{-n\beta}$ so that it remains to show that the second difference is exponentially small as well. We have

$$\begin{aligned}
&\sum_{z^n\in\mathcal{Z}^n} \big|\vartheta_{s,m_0}(z^n) - P_{Z_s^n|M_0}(z^n|m_0)\big| \\
&= \sum_{z^n\in\mathcal{Z}^n} \bigg|\frac{1}{|\mathcal{M}_1|}\sum_{m_1\in\mathcal{M}_1} \vartheta_{s,m_0}(z^n) \\
&\qquad - \frac{1}{|\mathcal{M}_1|}\sum_{m_1\in\mathcal{M}_1} \overline{W}^n_{\mathcal{Z},s}(z^n|m_0,m_1)\bigg| \\
&\leq \frac{1}{|\mathcal{M}_1|}\sum_{m_1\in\mathcal{M}_1}\sum_{z^n\in\mathcal{Z}^n} \big|\vartheta_{s,m_0}(z^n) - \overline{W}^n_{\mathcal{Z},s}(z^n|m_0,m_1)\big| \\
&\leq 2^{-n\beta}
\end{aligned}$$

so that

$$\|P_{Z_s^n M_1|M_0=m_0} - P_{Z_s^n|M_0=m_0} P_{M_1|M_0=m_0}\| \leq \lambda_n \tag{17}$$

with $\lambda_n = 2 \cdot 2^{-n\beta}$ for all $s \in \mathcal{S}$ and $m_0 \in \mathcal{M}_0$. Similarly as in Proposition 1, the continuity of the entropy function and $I(M_1; Z_s^n|M_0) = \frac{1}{|\mathcal{M}_0|}\sum_{m_0\in\mathcal{M}_0} I(M_1; Z_s^n|M_0 = m_0)$ imply now that

$$I(M_1; Z_s^n|M_0) \leq 2^{-n\beta/2} =: \epsilon_n$$

for n large enough proving the strong secrecy criterion (15).

To prove the second assertion (16), we write the average probability of decoding error at the non-legitimate receiver as

$$\begin{aligned}
\bar{e}'_{2,n}(s) &= \frac{1}{|\mathcal{M}_0||\mathcal{M}_1|}\sum_{m_0\in\mathcal{M}_0}\sum_{m_1\in\mathcal{M}_1} \overline{W}^n_{\mathcal{Z},s}(\widetilde{\mathcal{D}}^c_{s,m_0}(m_1)|m_0,m_1) \\
&= \frac{1}{|\mathcal{M}_0|}\sum_{m_0\in\mathcal{M}_0}\sum_{m_1\in\mathcal{M}_1} P_{Z_s^n M_1|M_0}\big(\widetilde{\mathcal{D}}^c_{s,m_0}(m_1), m_1|m_0\big) \\
&= \frac{1}{|\mathcal{M}_0|}\sum_{m_0\in\mathcal{M}_0} P_{Z_s^n M_1|M_0}\bigg(\bigcup_{m_1\in\mathcal{M}_1}\big\{\widetilde{\mathcal{D}}^c_{s,m_0}(m_1), m_1\big\}|m_0\bigg).
\end{aligned} \tag{18}$$

From the triangle inequality and from (17) we know that for all $s \in \mathcal{S}$ and $m_0 \in \mathcal{M}_0$ we have

$$\begin{aligned}
&\|P_{Z_s^n|M_0=m_0} P_{M_1|M_0=m_0}\| \\
&= \|P_{Z_s^n|M_0=m_0} P_{M_1|M_0=m_0} - P_{Z_s^n M_1|M_0=m_0} + P_{Z_s^n M_1|M_0=m_0}\| \\
&\leq \|P_{Z_s^n|M_0=m_0} P_{M_1|M_0=m_0} - P_{Z_s^n M_1|M_0=m_0}\| + \|P_{Z_s^n M_1|M_0=m_0}\| \\
&\leq \lambda_n + \|P_{Z_s^n M_1|M_0=m_0}\|
\end{aligned}$$

so that

$$\|P_{Z_s^n M_1|M_0=m_0}\| \geq \|P_{Z_s^n|M_0=m_0} P_{M_1|M_0=m_0}\| - \lambda_n$$

with $\lambda_n \to 0$ as $n \to \infty$. With this we can bound $\bar{e}'_{2,n}(s)$ in (18) from below by

$$\begin{aligned}
\bar{e}'_{2,n}(s) &\geq \frac{1}{|\mathcal{M}_0|}\sum_{m_0\in\mathcal{M}_0} \qquad\qquad (19) \\
&\quad\times P_{Z_s^n|M_0} P_{M_1|M_0}\bigg(\bigcup_{m_1\in\mathcal{M}_1}\big\{\widetilde{\mathcal{D}}^c_{s,m_0}(m_1), m_1\big\}|m_0\bigg) - \lambda_n \\
&= \frac{1}{|\mathcal{M}_0|}\sum_{m_0\in\mathcal{M}_0}\sum_{m_1\in\mathcal{M}_1} \\
&\quad\times P_{Z_s^n|M_0} P_{M_1|M_0}\big(\widetilde{\mathcal{D}}^c_{s,m_0}(m_1), m_1|m_0\big) - \lambda_n \\
&= \frac{1}{|\mathcal{M}_0||\mathcal{M}_1|}\sum_{m_0\in\mathcal{M}_0} \\
&\quad\times \sum_{m_1\in\mathcal{M}_1} P_{Z_s^n|M_0}(\widetilde{\mathcal{D}}^c_{s,m_0}(m_1)|m_0) - \lambda_n \\
&= \frac{1}{|\mathcal{M}_0||\mathcal{M}_1|}\sum_{m_0\in\mathcal{M}_0} \\
&\quad\times \sum_{m_1\in\mathcal{M}_1} \big(1 - P_{Z_s^n|M_0}(\widetilde{\mathcal{D}}_{s,m_0}(m_1)|m_0)\big) - \lambda_n \\
&= \frac{1}{|\mathcal{M}_0||\mathcal{M}_1|}\sum_{m_0\in\mathcal{M}_0} \big(|\mathcal{M}_1| - 1\big) - \lambda_n \\
&= 1 - \frac{1}{|\mathcal{M}_1|} - \lambda_n
\end{aligned}$$

for all $s \in \mathcal{S}$. Note that in the third step we used the fact that M_0 and M_1 are independent. This proves the second assertion (16). ■

The proposition shows that the non-legitimate receiver has the worst behavior of decoding performance as the decoding performance is the same as if he ignores his received signal $z^n \in \mathcal{Z}^n$ and simply selects a message $m_1 \in \mathcal{M}_1$ uniformly at random. In this case, the probability of success is $1/|\mathcal{M}_1|$ which is exactly expressed by (16).

But since $|\mathcal{M}_1| \to \infty$ and $\lambda_n \to 0$ exponentially fast as $n \to \infty$, the decoding error of the non-legitimate receiver in (16) approaches 1 exponentially fast meaning that he cannot decode the confidential message.

Finally, we collect all the properties and implications derived so far. As the vanishing output variation property guarantees strong secrecy in the sense of (6) and (15) we are able to generalize and extend the criterion as done in the following theorem.

Theorem 1: If a code for the compound BCC has the vanishing output variation property, i.e., $\|\overline{W}^n_{\mathcal{Z},s}(\cdot|m_0,m_1) - \vartheta_{s,m_0}\| \leq 2^{-n\beta}$ is satisfied for all $s \in \mathcal{S}$ and $(m_0,m_1) \in \mathcal{M}_0 \times \mathcal{M}_1$, then secrecy is guaranteed in the information theoretic sense of

$$\max_{s\in\mathcal{S}} \max\left\{I(M_1;Z_s^n), I(M_1;Z_s^n|M_0)\right\} \leq \epsilon_n$$

but also in the signal processing sense of

$$\min_{s\in\mathcal{S}} \bar{e}'_{2,n}(s) \geq 1 - \frac{1}{|\mathcal{M}_1|} - \lambda_n$$

with $1/|\mathcal{M}_1| \to 0$, $\epsilon_n \to 0$, and $\lambda_n \to 0$ exponentially fast as $n \to \infty$.

Remark 6: As Theorem 1 holds for any decoding strategy, there are no restrictions on the complexity or computational resources. This leads to universal results which hold for any applied post-processing strategy of the non-legitimate receiver.

C. Operational Meaning

So far we studied codes with vanishing output variation and showed that such codes realize secrecy in the sense of information theoretic strong secrecy, cf. (6) and (15), but also in the sense of the signal processing inspired approach of worst decoding performance, cf. (16).

Now, we directly connect the information theoretic criterion with the signal processing inspired concept. The following shows that strong secrecy $\max_{s\in\mathcal{S}} I(M_1;Z_s^n|M_0) \to 0$ exponentially fast implies worst behavior of decoding performance at the non-legitimate receiver, i.e., $\min_{s\in\mathcal{S}} \bar{e}'_{2,n}(s) \to 1$ exponentially fast. This gives the strong secrecy criterion an important signal processing meaning. This is particularly important as this result shows that any code (not necessarily having the vanishing output variation property) that realizes strong secrecy guarantees that average decoding error at the non-legitimate receiver goes to 1.

Corollary 1: The validity of the strong secrecy criterion $\max_{s\in\mathcal{S}} I(M_1;Z_s^n|M_0) \leq \epsilon_n$ immediately implies worst decoding performance, i.e., $\min_{s\in\mathcal{S}} \bar{e}'_{2,n}(s) \geq 1 - 1/|\mathcal{M}_1| - \lambda_n$.

Proof: The result follows immediately from Pinsker's inequality, cf. for example [24, Problem 3.18], and from previous Proposition 2. In more detail, for each $s \in \mathcal{S}$ we have

$$\begin{aligned}\epsilon_n &\geq I(M_1;Z_s^n|M_0)\\ &= \frac{1}{|\mathcal{M}_0|}\sum_{m_0\in\mathcal{M}_0} I(M_1;Z_s^n|M_0=m_0)\\ &= \frac{1}{|\mathcal{M}_0|}\sum_{m_0\in\mathcal{M}_0} D(P_{Z_s^nM_1|M_0=m_0}\|P_{Z_s^n|M_0=m_0}P_{M_1|M_0=m_0})\end{aligned}$$

$$\begin{aligned}&\geq \frac{1}{|\mathcal{M}_0|}\sum_{m_0\in\mathcal{M}_0}\frac{1}{2\ln 2}\big\|P_{Z_s^n|M_0=m_0}P_{M_1|M_0=m_0}\\ &\qquad\qquad - P_{Z_s^nM_1|M_0=m_0}\big\|^2\\ &\geq \frac{1}{|\mathcal{M}_0|}\sum_{m_0\in\mathcal{M}_0}\frac{1}{\ln 2}\\ &\quad\times\bigg(P_{Z_s^n|M_0}P_{M_1|M_0}\Big(\bigcup_{m_1\in\mathcal{M}_1}\{\widetilde{\mathcal{D}}^c_{s,m_0}(m_1),m_1\}|m_0\Big)\\ &\qquad -P_{Z_s^nM_1|M_0}\Big(\bigcup_{m_1\in\mathcal{M}_1}\{\widetilde{\mathcal{D}}^c_{s,m_0}(m_1),m_1\}|m_0\Big)\bigg)^2\end{aligned}$$

where the second last step follows from Pinsker's inequality and the last step from the definition of total variation distance. In particular, recall the definition $\|\mu - \nu\| := 2\sup_{A\subseteq\mathcal{A}}|\mu(A) - \nu(A)|$, cf. for example [22, Lemma 4.1.1], so that the last inequality is true for any sets, i.e., especially for the choice of $\bigcup_{m_1\in\mathcal{M}_1}\{\widetilde{\mathcal{D}}^c_{s,m_0}(m_1),m_1\}$. Using Jensen's inequality, we then obtain

$$\begin{aligned}\epsilon_n \geq \frac{1}{\ln 2}\bigg(&\frac{1}{|\mathcal{M}_0|}\sum_{m_0\in\mathcal{M}_0}\\ &\times P_{Z_s^n|M_0}P_{M_1|M_0}\Big(\bigcup_{m_1\in\mathcal{M}_1}\{\widetilde{\mathcal{D}}^c_{s,m_0}(m_1),m_1\}|m_0\Big)\\ &-\frac{1}{|\mathcal{M}_0|}\sum_{m_0\in\mathcal{M}_0}\\ &\times P_{Z_s^nM_1|M_0}\Big(\bigcup_{m_1\in\mathcal{M}_1}\{\widetilde{\mathcal{D}}^c_{s,m_0}(m_1),m_1\}|m_0\Big)\bigg)^2.\end{aligned}$$

Now observe that the second term corresponds to the formulation of the average probability of error given in (18). Thus, extracting the root and inserting this into (19) yields

$$\bar{e}'_{2,n}(s) \geq 1 - \frac{1}{|\mathcal{M}_1|} - \lambda_n$$

with $\lambda_n = \sqrt{\epsilon_n \ln 2}$ for all $s \in \mathcal{S}$ similarly as in Proposition 2 from (19) onwards. ■

IV. Achievable Secrecy Rate Region

The code design for the compound BCC is a challenging task. On the one hand, it has to protect the confidential message against the non-legitimate receiver 2 requiring a code which reveals no information to him. On the other hand, receiver 2 is also a legitimate receiver for the common message requiring a code suitable for reliable communication. At a first glance, these two intentions seem to be conflictive.

The previous section showed that a code having the vanishing output variation property is desirable for the secrecy task as it realizes secrecy in the information theoretic sense of the generalized criterion of strong secrecy, i.e,

$$\max_{s\in\mathcal{S}} \max\left\{I(M_1;Z_s^n), I(M_1;Z_s^n|M_0)\right\} \leq \epsilon_n \tag{20}$$

but also in the signal processing sense of worst decoding performance, i.e.,

$$\min_{s\in\mathcal{S}} \bar{e}'_{2,n}(s) \geq 1 - \frac{1}{|\mathcal{M}_1|} - \lambda_n, \tag{21}$$

cf. also Theorem 1. Next, we show that such codes can simultaneously be good for transmitting the common message to both receivers.

In the following we establish an achievable secrecy rate region for the compound BCC. As we make use of codes having vanishing output variation, we note that this result is valid for all discussed information theoretic and signal processing secrecy criteria. Therefore we simply refer to it as secrecy rate region and do not always explicitly state all the satisfied secrecy criteria.

Theorem 2: An achievable secrecy rate region for the compound BCC is given by the set of all rate pairs $(R_0, R_1) \in \mathbb{R}_+^2$ that satisfy

$$R_0 \leq \min_{s\in\mathcal{S}} \min\big\{I(U;Y_s), I(U;Z_s)\big\} \tag{22a}$$

$$R_1 \leq \min_{s\in\mathcal{S}} I(V;Y_s|U) - \max_{s\in\mathcal{S}} I(V;Z_s|U) \tag{22b}$$

for random variables $U - V - X - (Y_s, Z_s)$. Furthermore, the generalized criterion of strong secrecy (20) goes exponentially fast to zero and the decoding error (21) the non-legitimate receiver exponentially fast to one.

Proof: Instead of directly proving the achievability of (22), we will drop the auxiliary random variable V for a moment and will only show the achievability of

$$R_0 \leq \min_{s\in\mathcal{S}} \min\big\{I(U;Y_s), I(U;Z_s)\big\} \tag{23a}$$

$$R_1 \leq \min_{s\in\mathcal{S}} I(X;Y_s|U) - \max_{s\in\mathcal{S}} I(X;Z_s|U) \tag{23b}$$

for random variables $U - X - (Y_s, Z_s)$. Then this result generalizes to the whole region (22) including the auxiliary V by the following reasoning. Prefixing an artificial channel $P_{X|V} : \mathcal{V} \to \mathcal{P}(\mathcal{X})$ with finite $\mathcal{V}$ to the original channel $W_s = (W_{\mathcal{Y},s}, W_{\mathcal{Z},s})$ yields a "new" channel

$$\widetilde{W}_s(y,z|v) := \sum_{x\in\mathcal{X}} W_s(y,z|x) P_{X|V}(x|v) \tag{24}$$

which includes additional randomization. Clearly, the whole construction which we will carry out for W_s to prove (23), can immediately be applied to $\widetilde{W}_s$ to obtain a proof for (22).

In particular, the additional randomization will preserve the vanishing output variation property of a code. Recall that the corresponding definition already takes randomization (from the stochastic encoder) into account by a proper definition of the channel $\overline{W}^n_{\mathcal{Z},s} : \mathcal{M} \to \mathcal{P}(\mathcal{Z}^n)$, cf. (4). Thus, additional randomization as in (24) is easily incorporated by defining the channel as $\overline{W}^n_{\mathcal{Z},s}$ as

$$\overline{W}^n_{\mathcal{Z},s}(z^n|m_0,m_1) = \sum_{v^n\in\mathcal{V}^n} \widetilde{W}^n_{\mathcal{Z},s}(z^n|v^n) E(v^n|m_0,m_1).$$

We now come to the random coding proof of (23). For this we have to construct a codebook that realizes two tasks simultaneously: reliable communication of all messages to their respective receivers according to the rates given in (23) and secrecy of the confidential message. The constructed code will possess the vanishing output variation property, cf. Definition 4, so that from Theorem 1 we know that all discussed secrecy criteria will be satisfied.

In the following we extensively make use of the concept of *δ-typical* sequences from Csiszár and Körner [24] which is briefly recalled. Let $\delta > 0$. For any distribution $P_U \in \mathcal{P}(\mathcal{U})$, a sequence $u^n \in \mathcal{U}^n$ is called *δ-typical* if $|\frac{1}{n}N(u|u^n) - P_U(u)| \leq \delta$ for all $u \in \mathcal{U}$ and, in addition, $N(u|u^n) = 0$ if $P_U(u) = 0$. Here, $N(u|u^n)$ denotes the number of indices i such that $u_i = u$, $i = 1, ..., n$. The set of all such typical sequences is denoted by $\mathcal{T}^n_{U,\delta}$. Further, for any stochastic matrix $P_{X|U} : \mathcal{U} \to \mathcal{P}(\mathcal{X})$, a sequence $x^n \in \mathcal{X}^n$ is called *δ-typical* for given $u^n \in \mathcal{U}^n$ if $|\frac{1}{n}N(u,x|u^n,x^n) - \frac{1}{n}N(u|u^n)P_{X|U}(x|u)| \leq \delta$ for all $x \in \mathcal{X}$ and, in addition, $N(u,x|u^n,x^n) = 0$ if $P_{X|U}(x|u) = 0$. The set of all such sequences is denoted by $\mathcal{T}^n_{X|U,\delta}(u^n)$.

For probability distribution $P_U \in \mathcal{P}(\mathcal{U})$ and $\delta > 0$, we define the probability measure $P'_{U^n} \in \mathcal{P}(\mathcal{U}^n)$ as

$$P'_{U^n}(u^n) := \frac{P^n_U(u^n)}{P^n_U(\mathcal{T}^n_{U,\delta})} \tag{25}$$

if $u^n \in \mathcal{T}^n_{U,\delta}$ and $P'_{U^n}(u^n) = 0$ else, where $P^n_U(u^n) = \prod_{i=1}^n P_U(u_i)$. Similarly, for $P_{X|U} : \mathcal{U} \to \mathcal{P}(\mathcal{X})$ we define $P'_{X^n|U^n} : \mathcal{U}^n \to \mathcal{P}(\mathcal{X}^n)$ as

$$P'_{X^n|U^n}(x^n|u^n) := \frac{P^n_{X|U}(x^n|u^n)}{P^n_{X|U}(\mathcal{T}^n_{X|U,\delta}(u^n)|u^n)} \tag{26}$$

if $x^n \in \mathcal{T}^n_{X|U,\delta}(u^n)$ and $P'_{X^n|U^n}(x^n|u^n) = 0$ else, where $P^n_{X|U}(x^n|u^n) = \prod_{i=1}^n P_{X|U}(x_i|u_i)$.

Let $\mathcal{M}_0$ be the set of common messages with size $M_{0,n} = \lfloor \min_{s\in\mathcal{S}}\{2^{n(I(U;Y_s)-\tau/2}, 2^{n(I(U;Z_s)-\tau/2}\}\rfloor$. Let $\mathcal{M}_1$ be the set of confidential messages and further $\mathcal{L} := \{1, ..., L_n\}$ satisfying $M'_{1,n} := L_n M_{1,n} = \lfloor \min_{s\in\mathcal{S}} 2^{n(I(X;Y_s|U)-\tau/2)}\rfloor$ for some (small) $\tau > 0$. Thereby, the set $\mathcal{L}$ will carry no confidential information, but "dummy" messages for additional randomization. The actual sizes of L_n and $M_{1,n}$ will be determined later and the main question is how many resources have to be allocated to that "dummy" message set in order to ensure secrecy of the confidential messages.

First, generate $M_{0,n}$ independent random codewords $U^n_{m_0} \in \mathcal{U}^n$ with $m_0 \in \mathcal{M}_0$ according to P'_{U^n}, cf. (25). Then, for each $U^n_{m_0} \in \mathcal{U}^n$ generate $L_n M_{1,n}$ independent random codewords $X^n_{lm_1m_0} \in \mathcal{X}^n$ with $l \in \mathcal{L}$ and $m_1 \in \mathcal{M}_1$ according to $P'_{X^n|U^n}$, cf. (26).

Let us drop the secrecy requirement for a moment and interpret $m'_1 = (l, m_1) \in \mathcal{L} \times \mathcal{M}_1 = \mathcal{M}'_1$ as a public message intended for receiver 1 which need not be kept secret from non-legitimate receiver 2. Then this scenario is related to the broadcast channel with degraded message sets [26].

Lemma 1: With the random coding scheme defined above, all rate pairs $(R_0, R'_1) \in \mathbb{R}_+^2$ that satisfy

$$R_0 \leq \min_{s\in\mathcal{S}} \min\big\{I(U;Y_s), I(U;Z_s)\big\} \tag{27a}$$

$$R'_1 \leq \min_{s\in\mathcal{S}} I(X;Y_s|U) \tag{27b}$$

for random variables $U - X - (Y_s, Z_s)$ are achievable for the compound broadcast channel with degraded message sets. In particular, these rates are achievable with average probability of errors of the form $\bar{e}_{1,n}, \bar{e}_{2,n} \leq 2^{-n\gamma}$ for some $\gamma > 0$.

Sketch of Proof: A random codebook as defined above, i.e., a superposition of codewords for the common message and for the public message according to the chosen input distributions (25) and (26), will allow to prove the result in a similar way as for example in [27] for the compound bidirectional broadcast channel. The details are omitted. ∎

Remark 7: For $|\mathcal{S}| = 1$ the region (27) reduces to a subregion of [26]. More precisely, the sum constraint on receiver 1 of the form $R_0 + R_1 \leq I(X; Y_s)$ in [26] is replaced by individual constraints on $R_0 \leq I(U; Y_s)$ and $R_1 \leq I(X; Y_s|U)$ which makes the region smaller. However, (27) will be sufficient to establish the desired result in (22).

Next we incorporate the secrecy requirement on the confidential message, where we want to exploit the concept of vanishing output variation, cf. Definition 4 and Theorem 1. For this purpose, we have to carefully choose the desired channel outputs at the non-legitimate receiver 2 (given by the measures ϑ_{s,m_0}, $s \in \mathcal{S}$, $m_0 \in \mathcal{M}_0$). The set $\mathcal{L}$ of "dummy" messages will then be used to ensure that the code will possess the vanishing output variation property such that (5) is satisfied (which then implies secrecy, cf. Theorem 1). Accordingly, the main important points to address in the following are: first, how should the measures ϑ_{s,m_0} be chosen and, second, how large should be the size of $\mathcal{L}$.

To address the first point, we note that the channel $W_{\mathcal{Z},s}$ can also be regarded as a channel with inputs in $\mathcal{U} \times \mathcal{X}$ where the $\mathcal{U}$-inputs do not make any difference. Moreover, it will be sufficient to concentrate only on those outputs that are typical; the probability of all other outputs will be of no consequence as we will see later. Therefore, we define for every channel realization $s \in \mathcal{S}$, message triple $(l, m_1, m_0) \in \mathcal{L} \times \mathcal{M}_1 \times \mathcal{M}_0$, and channel output $z^n \in \mathcal{Z}^n$ the random variable

$$\begin{aligned} &Q_s^n(z^n|X^n_{lm_1m_0}, U^n_{m_0}) \\ &\quad := W^n_{\mathcal{Z},s}(z^n|X^n_{lm_1m_0})\mathbb{1}_{\mathcal{T}^n_{Z_s|XU,\delta}(X^n_{lm_1m_0},U^n_{m_0})}(z^n), \end{aligned} \tag{28}$$

where for any set $\mathcal{A} \subset \mathcal{Z}^n$, we let $\mathbb{1}_{\mathcal{A}}(z^n) = 1$ if $z^n \in \mathcal{A}$ and $\mathbb{1}_{\mathcal{A}}(z^n) = 0$ else. Conditional on $U^n_{m_0}$, these random variables are i.i.d. Moreover, as the input $(X^n_{lm_1m_0}, U^n_{m_0})$ is jointly 2δ-typical with respect to the joint distribution P_{XU}, and the outputs of Q_s^n are δ-typical conditional on the inputs, we know from [24] that (28) is bounded from above by

$$Q_s^n(z^n|X^n_{lm_1m_0}, U^n_{m_0}) \leq 2^{-n(H(Z_s|X,U)-\delta_1)} \tag{29}$$

for some $\delta_1 = \delta_1(\delta)$. Now let

$$\vartheta'_{s,U^n_{m_0}}(z^n) = \mathbb{E}\big[Q_s^n(z^n|X^n_{lm_1m_0}, U^n_{m_0})|U^n_{m_0}\big]$$

be the expectation of (28) conditional on $U^n_{m_0}$. Note that due to construction, $\vartheta'_{s,U^n_{m_0}}$ is a non-negative measure. For any $\epsilon_n > 0$ we define

$$\begin{aligned} \mathcal{F}_{s,U^n_{m_0}} := \big\{&z^n \in \mathcal{T}^n_{Z_s|U,2|\mathcal{X}|\delta}(U^n_{m_0}) : \\ &\vartheta'_{s,U^n_{m_0}}(z^n) \geq \epsilon_n |\mathcal{T}^n_{Z_s|U,2|\mathcal{X}|\delta}(U^n_{m_0})|^{-1}\big\}. \end{aligned} \tag{30}$$

Finally, we set

$$\vartheta_{s,U^n_{m_0}}(z^n) := \vartheta'_{s,U^n_{m_0}}(z^n)\mathbb{1}_{\mathcal{F}_{s,U^n_{m_0}}}(z^n). \tag{31}$$

As we will see later in the proof, this will be a suitable choice for the desired channel output at the non-legitimate receiver. Next, we turn our attention to the second important point, which is the size of $\mathcal{L}$. Similarly, we set

$$\widetilde{Q}_s^n(z^n|X^n_{lm_1m_0}, U^n_{m_0}) = Q_s^n(z^n|X^n_{lm_1m_0}, U^n_{m_0})\mathbb{1}_{\mathcal{F}_{s,U^n_{m_0}}}(z^n).$$

Then we define the event $\mathcal{Q}_{s,U^n_{m_0}}(z^n)$ as

$$\frac{1}{L_n}\sum_{l=1}^{L_n} \widetilde{Q}_s^n(z^n|X^n_{lm_1m_0}, U^n_{m_0}) \in \big[(1 \pm \epsilon_n)\vartheta_{s,U^n_{m_0}}(z^n)\big]. \tag{32}$$

For the analysis of this event we need a bound on the concentration of sums of i.i.d. random variables around their expectation as given in the following lemma which is due to Chernoff and Hoeffding.

Lemma 2: Let $b > 0$ and $Z_1, Z_2, ..., Z_L$ be i.i.d. random variables with values in $[0, b]$. Further, let $\mu = \mathbb{E}[Z_1]$ be the expectation of Z_1. Then

$$\mathbb{P}\left\{\frac{1}{L}\sum_{l=1}^{L} Z_l \notin [(1 \pm \epsilon)\mu]\right\} \leq 2\exp\left(-L \cdot \frac{\epsilon^2\mu}{2b\ln 2}\right)$$

where $[(1 \pm \epsilon)\mu]$ denotes the interval $[(1-\epsilon)\mu, (1+\epsilon)\mu]$.

Proof: A proof can be found in [28] or [29]. ∎

Now let $z^n \in \mathcal{Z}^n$. Then the probability of the complement of $\mathcal{Q}_{s,U^n_{m_0}}(z^n)$ is

$$\begin{aligned} &\mathbb{P}\big\{(\mathcal{Q}_{s,U^n_{m_0}}(z^n))^c\big\} \\ &\quad= \sum_{u^n \in \mathcal{U}^n} \mathbb{P}\{U^n_{m_0} = u^n\}\mathbb{P}\big\{(\mathcal{Q}_{s,U^n_{m_0}}(z^n))^c|U^n_{m_0} = u^n\big\} \\ &\quad\leq \sum_{u^n \in \mathcal{U}^n} \mathbb{P}\{U^n_{m_0} = u^n\} \\ &\qquad\quad \times 2\exp\left(-L_n \cdot \frac{\epsilon_n^2 2^{n(H(Z_s|X,U)-\delta_1)}\vartheta_{s,u^n}(z^n)}{2\ln 2}\right) \\ &\quad\leq 2\exp\left(-L_n \cdot \frac{\epsilon_n^3 2^{-n(I(X;Z_s|U)+\delta_1+\delta_2)}}{2\ln 2}\right) \end{aligned} \tag{33}$$

where the first step follows from the law of total probability, the second step from Lemma 2 (with $\vartheta_{s,U^n_{m_0}}(z^n)$ in the role of μ) and (29), and the last step from (30) and

$$\big|\mathcal{T}^n_{Z_s|U,2|\mathcal{X}|\delta}(U^n_{m_0})\big| \leq 2^{n(H(Z_s|U)+\delta_2)}$$

for some $\delta_2 = \delta_2(\delta)$, see [24], which applies here since $U^n_{m_0}$ is δ-typical. Note that if we choose $\epsilon_n = 2^{-n\beta}$ for some $\beta \leq \frac{1}{4}\min\{\gamma, \delta_1 + \delta_2\}$, then (33) tends to zero doubly-exponentially fast for

$$L_n \geq 2^{n(\max_{s \in \mathcal{S}} I(X;Z_s|U)+2(\delta_1+\delta_2))}. \tag{34}$$

Note that we have to choose the maximum in (34) to ensure that (33) tends to zero doubly-exponentially for all channel realizations $s \in \mathcal{S}$.

Next, we determine the sizes of the remaining sets for the confidential message. For $\max_{s \in \mathcal{S}} I(X; Z_s|U) < \min_{s \in \mathcal{S}} I(X; Y_s|U)$, we choose δ (and therewith also

δ_1 and δ_2) small enough such that (34) is satisfied and at the same time

$$L_n \leq 2^{n(\max_{s\in\mathcal{S}} I(X;Z_s|U)+3(\delta_1+\delta_2))} \tag{35a}$$
$$< 2^{n(\min_{s\in\mathcal{S}} I(X;Y_s|U)-\tau/2)}. \tag{35b}$$

With (34) and (35) we have determined the size of the "dummy" message set $\mathcal{L}$. This provides the basis to prove that this code has the vanishing output variation property, cf. (5). Once this property is guaranteed, the discussion in Section III ensures the secrecy of the confidential message. Further, for the confidential messages we set

$$M_{1,n} \leq 2^{n(\min_{s\in\mathcal{S}} I(X;Y_s|U)-\tau/2-\max_{s\in\mathcal{S}} I(X;Z_s|U)-3(\delta_1+\delta_2))}.$$

From (33)-(34) we know that (32) is satisfied for every $s \in \mathcal{S}$, $(m_0, m_1) \in \mathcal{M}_0 \times \mathcal{M}_1$, and $z^n \in \mathcal{Z}^n$ with probability close to one. Further, with $\mathcal{M}_1' = \mathcal{L} \times \mathcal{M}_1$ we know from Lemma 1 that the random codewords we have chosen are the codewords of a deterministic code achieving $\bar{e}_{1,n}, \bar{e}_{2,n} \leq 2^{-n\gamma}$ for some $\gamma > 0$ with probability close to one. Thus, there must be realizations of $(U_{m_0}^n, X_{lm_1m_0}^n)$ and $\vartheta_{s,U_{m_0}^n}$ with both these properties, which we denote by $(u_{m_0}^n, x_{lm_1m_0}^n)$ and ϑ_{s,m_0} respectively.

From this we obtain an appropriate code with a stochastic encoder as follows. Each message pair $(m_0, m_1) \in \mathcal{M}_0 \times \mathcal{M}_1$ is mapped into the codeword $x_{lm_1m_0}^n \in \mathcal{X}^n$ with probability $1/L_n$ which defines a stochastic encoder as given in Definition 2, cf. also Remark 2. The decoder at legitimate receiver 1 decodes all indices, i.e., (l, m_1, m_0), while the decoder at non-legitimate receiver 2 only decodes the common message m_0. Interpreting (l, m_1) as a public message m_1' for receiver 1, we know from Lemma 1 that this code is suitable for reliable transmission of all messages to their respective receivers. It remains to prove that this code has also the vanishing output variation property, cf. (5).

From the triangle inequality we obtain for every $s \in \mathcal{S}$ and $(m_0, m_1) \in \mathcal{M}_0 \times \mathcal{M}_1$

$$\begin{aligned}
&\big\|\overline{W}_{\mathcal{Z},s}^n(\cdot|m_0, m_1) - \vartheta_{s,m_0}\big\| \\
&\quad\leq \Big\|\overline{W}_{\mathcal{Z},s}^n(\cdot|m_0, m_1) - \frac{1}{L_n}\sum_{l=1}^{L_n} Q_s^n(\cdot|x_{lm_1m_0}^n, u_{m_0}^n)\Big\| \\
&\qquad+\Big\|\frac{1}{L_n}\sum_{l=1}^{L_n} Q_s^n(\cdot|x_{lm_1m_0}^n, u_{m_0}^n)(1-\mathbb{1}_{\mathcal{F}_{s,m_0}})\Big\| \\
&\qquad+\Big\|\frac{1}{L_n}\sum_{l=1}^{L_n} Q_s^n(\cdot|x_{lm_1m_0}^n, u_{m_0}^n)\mathbb{1}_{\mathcal{F}_{s,m_0}} - \vartheta_{s,m_0}\Big\|.
\end{aligned}$$

We denote the three parts by I, II, and III in that order and bound each of them separately. Since all codewords satisfy (32), we immediately have for the third term $III \leq \epsilon$.

For the first term I we have

$$\begin{aligned}
I &\leq \frac{1}{L_n}\sum_{l=1}^{L_n}\Big\|W_{\mathcal{Z},s}^n(\cdot|x_{lm_1m_0}^n, u_{m_0}^n) - Q_s^n(\cdot|x_{lm_1m_0}^n, u_{m_0}^n)\Big\| \\
&= \frac{1}{L_n}\sum_{l=1}^{L_n}\Big\|W_{\mathcal{Z},s}^n(\cdot|x_{lm_1m_0}^n, u_{m_0}^n)\big(1-\mathbb{1}_{\mathcal{T}_{Z_s|XU,\delta}^n(x_{lm_1m_0}^n, u_{m_0}^n)}\big)\Big\| \\
&= \frac{1}{L_n}\sum_{l=1}^{L_n} W_{\mathcal{Z},s}^n\big((\mathcal{T}_{Z_s|XU,\delta}^n(x_{lm_1m_0}^n, u_{m_0}^n))^c|x_{lm_1m_0}^n, u_{m_0}^n\big) \\
&\leq (n+1)^{|\mathcal{U}||\mathcal{X}||\mathcal{Z}|}2^{-nc\delta^2}
\end{aligned}$$

for some constant $c > 0$, where we again interpret $W_{\mathcal{Z},s}$ as a channel from $\mathcal{U} \times \mathcal{X}$ to $\mathcal{Z}$ and use the fact that the probability that the output of a channel is not δ-typical conditional on the inputs is exponentially small, cf. for example [27, Lemma 2] or [30, Lemma III.1.3].

Finally, the second term II can be rewritten as

$$\begin{aligned}
II &\leq 1 - \frac{1}{L_n}\sum_{l=1}^{L_n} Q_s^n(\mathcal{F}_{s,m_0}|x_{lm_1m_0}^n, u_{m_0}^n) \\
&\leq 1-(1-\epsilon_n)\vartheta_{s,m_0}(\mathcal{F}_{s,m_0}) \\
&= 1-(1-\epsilon_n)\vartheta_{s,m_0}'(\mathcal{F}_{s,m_0}) \\
&= 1-(1-\epsilon_n)\big(\vartheta_{s,m_0}'(\mathcal{T}_{Z_s|U,2|\mathcal{X}|\delta}^n(u_{m_0}^n)) \\
&\qquad\qquad - \vartheta_{s,m_0}'(\mathcal{T}_{Z_s|U,2|\mathcal{X}|\delta}^n(u_{m_0}^n)\backslash\mathcal{F}_{s,m_0})\big)
\end{aligned} \tag{36}$$

where the second step follows from (32) and the third from the fact that $\vartheta_{s,m_0}(\mathcal{F}_{s,m_0}) = \vartheta_{s,m_0}'(\mathcal{F}_{s,m_0})$, cf. (31), and the last step from the definition of $\mathcal{F}_{s,m_0}$, cf. (30). Now we have

$$\begin{aligned}
&\vartheta_{s,m_0}'(\mathcal{T}_{Z_s|U,2|\mathcal{X}|\delta}^n(u_{m_0}^n)) \\
&\quad= \mathbb{E}\big[Q_s^n(\mathcal{T}_{Z_s|U,2|\mathcal{X}|\delta}^n(U_{m_0}^n)|X_{lm_1m_0}^n, U_{m_0}^n)|u_{m_0}^n\big] \\
&\quad\geq \mathbb{E}\big[W_{\mathcal{Z},s}^n(\mathcal{T}_{Z_s|XU,\delta}^n(X_{lm_1m_0}^n, U_{m_0}^n)|X_{lm_1m_0}^n, U_{m_0}^n)|u_{m_0}^n\big] \\
&\quad\geq 1-(n+1)^{|\mathcal{U}||\mathcal{X}||\mathcal{Z}|}2^{-nc\delta^2}
\end{aligned} \tag{37}$$

and further

$$\vartheta_{s,m_0}'(\mathcal{T}_{Z_s|U,2|\mathcal{X}|\delta}^n(u_{m_0}^n)\backslash\mathcal{F}_{s,m_0}) \leq \epsilon_n. \tag{38}$$

Now, with (37) and (38) in (36), we obtain for the second term

$$\begin{aligned}
II &\leq 1-(1-\epsilon_n)(1-(n+1)^{|\mathcal{U}||\mathcal{X}||\mathcal{Z}|}2^{-nc\delta^2}-\epsilon_n) \\
&\leq 2\epsilon_n + (n+1)^{|\mathcal{U}||\mathcal{X}||\mathcal{Z}|}2^{-nc\delta^2}.
\end{aligned}$$

Putting all three terms together, we can bound the total variation distance as

$$\big\|\overline{W}_{\mathcal{Z},s}^n(\cdot|m_0, m_1)-\vartheta_{s,m_0}\big\| \leq 3\epsilon_n+2(n+1)^{|\mathcal{U}||\mathcal{X}||\mathcal{Z}|}2^{-nc\delta^2} \tag{39}$$

which proves (5). Note that (39) becomes exponentially small since we chose $\epsilon_n = 2^{-n\beta}$. Thus, our code has exponentially fast vanishing output variation according to Definition 4, cf. (5), so that the secrecy criteria discussed in Section III, cf. especially Theorem 1, are satisfied as well. ■

V. Outer Bound and Multi-Letter Description

Here we consider the counterpart of Theorem 2 and establish a multi-letter characterization of an outer bound on the strong secrecy capacity region. For this purpose we define the region $\mathcal{R}$ as the set of all rate pairs $(R_0, R_1) \in \mathbb{R}_+^2$ that satisfy

$$R_0 \leq \lim_{n\to\infty}\tfrac{1}{n}\inf_{s\in\mathcal{S}}\min\big\{I(U;Y_s^n), I(U;Z_s^n)\big\} \tag{40a}$$
$$R_1 \leq \lim_{n\to\infty}\tfrac{1}{n}\big(\inf_{s\in\mathcal{S}} I(V;Y_s^n|U) - \sup_{s\in\mathcal{S}} I(V;Z_s^n|U)\big) \tag{40b}$$

for random variables satisfying the Markov chain relationship $U - V - X^n - (Y_s^n, Z_s^n)$. This means, the joint probability distribution is specified by

$$\begin{aligned} &P_{UVX^nY_s^nZ_s^n}(u, v, x^n, y^n, z^n) \\ &\quad = W_s^n(y^n, z^n|x^n) P_{X^n|V}(x^n|v) P_{V|U}(v|u) P_U(u) \\ &\quad = \prod_{i=1}^{n} W_s(y_i, z_i|x_i) P_{X^n|V}(x^n|v) P_{V|U}(v|u) P_U(u). \quad (41)\end{aligned}$$

We further need the following lemma.

Lemma 3: Let $\mathfrak{W} := \{(W_{\mathcal{Y},s}, W_{\mathcal{Z},s}) : s \in \mathcal{S}\}$ be an arbitrary compound broadcast channel. For random variables $U - V - X^n - (Y_s^n, Z_s^n)$, the limit

$$\lim_{n\to\infty} \tfrac{1}{n}\Big(\inf_{s\in\mathcal{S}} I(V; Y_s^n|U) - \sup_{s\in\mathcal{S}} I(V; Z_s^n|U)\Big)$$

exists and is equal to $\sup_{n\in\mathbb{N}} \frac{1}{n}(\inf_{s\in\mathcal{S}} I(V; Y_s^n|U) - \sup_{s\in\mathcal{S}} I(V; Z_s^n|U))$.

Proof: We follow [15] and use Fekete's lemma [31] to prove the desired result. We have to show that the sequence $(a_n)_{n\in\mathbb{N}}$ with

$$a_n := \inf_{s\in\mathcal{S}} I(V; Y_s^n|U) - \sup_{s\in\mathcal{S}} I(V; Z_s^n|U)$$

satisfies

$$a_{n+m} \geq a_n + a_m$$

for all $n, m \in \mathbb{N}$. Therefore, we define Markov chains $U_1 - V_1 - X^n - (Y_s^n, Z_s^n)$ and $U_2 - V_2 - \widetilde{X}^m - (\widetilde{Y}_s^m, \widetilde{Z}_s^m)$ and set $U := (U_1, U_2)$, $V := (V_1, V_2)$, $X^{n+m} := (X^n, \widetilde{X}^m)$, and $(Y_s^{n+m}, Z_s^{n+m}) := ((Y_s^n, \widetilde{Y}_s^m), (Z_s^n, \widetilde{Z}_s^m))$. By the definition of a_n we have

$$\begin{aligned} a_{n+m} &= \inf_{s\in\mathcal{S}} I(V; Y_s^{n+m}|U) - \sup_{s\in\mathcal{S}} I(V; Z_s^{n+m}|U) \\ &\geq \inf_{s\in\mathcal{S}} I(V_1; Y_s^n|U_1) + \inf_{s\in\mathcal{S}} I(V_2; \widetilde{Y}_s^m|U_2) \\ &\quad - \sup_{s\in\mathcal{S}} I(V_1; Z_s^n|U_1) - \sup_{s\in\mathcal{S}} I(V_2; \widetilde{Z}_s^m|U_2)\end{aligned}$$

which follows from the independence of the two Markov chains. Since these Markov chains can be arbitrary, we conclude that $a_{n+m} \geq a_n + a_m$ holds for all $n, m \in \mathbb{N}$. ■

Theorem 3: The region $\mathcal{R}$ given in (40) is a multi-letter outer bound on the strong secrecy capacity region $\mathcal{C}_S(\mathfrak{W})$ of the compound BCC $\mathfrak{W}$, i.e., we have

$$\mathcal{C}_S(\mathfrak{W}) \subseteq \mathcal{R}.$$

Proof: For any given sequence of $(n, M_{0,n}, M_{1,n})$-codes of Definition 2 with $\bar{e}_{1,n}, \bar{e}_{2,n} \to 0$ and

$$\sup_{s\in\mathcal{S}} I(M_1; Z_s^n) = H(M_1) - \inf_{s\in\mathcal{S}} H(M_1|Z_s^n) =: \epsilon_{c,n} \quad (42)$$

with $\epsilon_{c,n} \to 0$, there exist random variables $U - V - X^n - (Y_s^n, Z_s^n)$ such that all rate pairs $(R_0, R_1) \in \mathbb{R}_+^2$ are bounded by (40).

Let M_0 and M_1 be random variables uniformly distributed over the message sets $\mathcal{M}_0$ and $\mathcal{M}_1$. We have the Markov chains

$$(M_0, M_1) - X^n - Y_s^n - (\hat{M}_{0,1}, \hat{M}_1)$$

and

$$(M_0, M_1) - X^n - Z_s^n - \hat{M}_{0,2}$$

where in both cases the first transition is governed by the stochastic encoder E, cf. (1), the second by the corresponding channel $W_{\mathcal{Y},s}^n$ and $W_{\mathcal{Z},s}^n$, and last one by the corresponding decoder, cf. (2a) and (2b). Then we have for all $s \in \mathcal{S}$ at receiver 1 for the common rate

$$\begin{aligned} nR_0 &= H(M_0) \\ &= I(M_0; Y_s^n) + H(M_0|Y_s^n) \\ &\leq I(M_0; Y_s^n) + n\epsilon_{1,n} \quad (43)\end{aligned}$$

where the last inequality follows from Fano's inequality, i.e., $H(M_0|Y_s^n) \leq H(M_0, M_1|Y_s^n) \leq n\epsilon_{1,n}$, and similarly for all $s \in \mathcal{S}$ at receiver 2 we get

$$nR_0 = H(M_0) \leq I(M_0; Z_s^n) + n\epsilon_{2,n} \quad (44)$$

by using Fano's inequality $H(M_0|Z_s^n) \leq n\epsilon_{2,n}$.

Next, we follow [5] and make use of the definition of mutual information. Rewriting (42) we get for the confidential rate

$$\begin{aligned} nR_1 &= H(M_1) \\ &= \inf_{s\in\mathcal{S}} H(M_1|Z_s^n) + \epsilon_{c,n} \\ &= \inf_{s\in\mathcal{S}} \big(H(M_1|Z_s^n, M_0) + I(M_1; M_0|Z_s^n)\big) + \epsilon_{c,n} \\ &\leq H(M_1|M_0) - \sup_{s\in\mathcal{S}} I(M_1; Z_s^n|M_0) + n\epsilon_{2,n} + \epsilon_{c,n} \\ &\leq I(M_1; Y_s^n|M_0) - \sup_{s'\in\mathcal{S}} I(M_1; Z_{s'}^n|M_0) + n\epsilon_n \quad (45)\end{aligned}$$

with $\epsilon_n = \epsilon_{1,n} + \epsilon_{2,n} + \epsilon_{c,n}$ where the first inequality follows from $I(M_1; M_0|Z_s^n) = H(M_0|Z_s^n) - H(M_0|Z_s^n, M_1) \leq H(M_0|Z_s^n) \leq n\epsilon_{2,n}$ and the second inequality from $H(M_1|Y_s^n, M_0) \leq H(M_0, M_1|Y_s^n) \leq n\epsilon_{1,n}$.

With $I(M_1; Y_s^n|M_0) = I(M_0, M_1; Y_s^n|M_0)$ and $I(M_1; Z_s^n|M_0) = I(M_0, M_1; Z_s^n|M_0)$, (43)-(45) imply that the rates are bounded by

$$\begin{aligned} nR_0 &\leq \inf_{s\in\mathcal{S}} \min\big\{I(M_0; Y_s^n), I(M_0; Z_s^n)\big\} + n\epsilon_n \\ nR_1 &\leq \inf_{s\in\mathcal{S}} I(M_0, M_1; Y_s^n|M_0) - \sup_{s\in\mathcal{S}} I(M_0, M_1; Z_s^n|M_0) + n\epsilon_n.\end{aligned}$$

Recall that the transition between the messages (M_0, M_1) and the input X^n is governed by a stochastic encoder E, cf. (1), which allows us to introduce arbitrary auxiliary random variables U and V which satisfy the Markov chain $U - V - X^n - (Y_s^n, Z_s^n)$. Dividing by n and taking the limit yields

$$\begin{aligned} R_0 &\leq \lim_{n\to\infty} \tfrac{1}{n} \inf_{s\in\mathcal{S}} \min\big\{I(U; Y_s^n), I(U; Z_s^n)\big\} \\ R_1 &\leq \lim_{n\to\infty} \tfrac{1}{n}\Big(\inf_{s\in\mathcal{S}} I(V; Y_s^n|U) - \sup_{s\in\mathcal{S}} I(V; Z_s^n|U)\Big)\end{aligned}$$

where Lemma 3 guarantees that the quantities on right hand side exist and are well defined. This concludes the proof. ■

The multi-letter outer bound given in Theorem 3 and the achievability result in Theorem 2 applied to the n-fold product of the broadcast channel yields the following.

Corollary 2: A multi-letter description of the strong secrecy capacity region $\mathcal{C}_S(\mathfrak{W})$ of the compound BCC $\mathfrak{W}$ is given by

$$\mathcal{C}_S(\mathfrak{W}) = \mathcal{R}.$$

Proof: Consider the n-fold product of the broadcast channel $W_s^n : \mathcal{X}^n \to \mathcal{P}(\mathcal{Y}^n \times \mathcal{Z}^n)$ and further the auxiliary channel $P_{X^n|V} : \mathcal{V} \to \mathcal{P}(\mathcal{X}^n)$, cf. also (41). Together this defines a channel from $\mathcal{V}$ to $\mathcal{Y}^n \times \mathcal{Z}^n$ specified by

$$\widehat{W}_s(y^n, z^n|v) := \sum_{x^n \in \mathcal{X}^n} W_s^n(y^n, z^n|x^n) P_{X^n|V}(x^n|v) \quad (46)$$

which collects the n channel uses into one new "*block*." Applying the one-shot achievability result given in Theorem 2 to the corresponding marginal channels $\widehat{W}_{\mathcal{Y},s}$ and $\widehat{W}_{\mathcal{Z},s}$ of the "*blocked*" channel in (46) yields $\mathcal{R}$ as an achievable rate region for the corresponding multi-letter case. Note that this blocking preserves the vanishing output variation property of the code, cf. Definition 4, so that the information theoretic and signal processing secrecy criteria (as discussed in Section III and Theorem 1) are satisfied.

The multi-letter outer bound given Theorem 3 yields the matching converse and completes the proof. ■

Corollary 3: For all rate pairs $(R_0, R_1) \in \mathcal{R}$ there are coding schemes that realize the generalized strong secrecy criterion (20), i.e., $\max_{s\in\mathcal{S}} \max\big\{I(M_1; Z_s^n), I(M_1; Z_s^n|M_0)\big\} \to 0$ exponentially fast, and also worst decoding performance of the non-legitimate receiver (21), i.e., $\min_{s\in\mathcal{S}} \bar{e}'_{2,n}(s) \to 1$ exponentially fast for all decoding strategies.

Proof: The result follows immediately from Theorem 1 and Corollary 2. ■

Remark 8: In the original definition of achievability, the strong secrecy criterion is defined in the traditional way as $\max_{s\in\mathcal{S}} I(M_1; Z_s^n) \leq \epsilon_n$, cf. Definition 3 and (3), which yields the secrecy capacity region $\mathcal{C}_S(\mathfrak{W})$. Obviously, one could state an own formal definition of achievability with the generalized secrecy criterion $\max_{s\in\mathcal{S}} \max\{I(M_1; Z_s^n), I(M_1; Z_s^n|M_0)\} \leq \epsilon_n$, cf. (20). However, from Corollary 2 we know that this would lead to the same secrecy capacity region as for the traditional definition. The reasoning is the following. Obviously, the generalized secrecy criterion (20) is stronger than the traditional one in (3). Thus, the corresponding secrecy region must be contained in $\mathcal{C}_S(\mathfrak{W})$. But Corollary 2 actually shows that we achieve $\mathcal{C}_S(\mathfrak{W})$ also for the generalized secrecy criterion, both regions must coincide.

VI. Conclusion

In this paper we studied robust broadcasting of common and confidential messages over compound channels. This can be modeled by the compound broadcast channel with confidential messages (BCC), where a transmitter sends a common message to two receivers and a confidential message intended for receiver 1 which has to be kept secret from receiver 2. Thus, receiver 2 is a legitimate receiver for the common message and, at the same time, a non-legitimate receiver for the confidential message. This necessitated a careful code design, which realizes both conflicting tasks simultaneously.

We questioned the validity of the classical definition of information theoretic strong secrecy in our scenario, since the non-legitimate receiver 2 is part of the communication system. As he is intended to decode the common message, this may be available as side information for inferring the confidential message. Accordingly we generalized the strong secrecy criterion taking such side information into account.

Along with this, we investigated the code concept of vanishing output variation and showed that such codes guarantee both notions of strong secrecy (the classical and the generalized version). Moreover, such codes also yield the worst decoding performance at the non-legitimate receiver regardless of his computational capabilities. This makes the concept of vanishing output variation particularly desirable, since it realizes secrecy from the information theoretic but also from the signal processing point of view.

We further derived an achievable secrecy rate region for the compound BCC and presented a multi-letter outer bound. Both together establishes a multi-letter characterization of the corresponding secrecy capacity region. Thereby, the secrecy capacity region reveals an interesting structure, since different assumptions have to be made on the channel to receiver 2. As he is a legitimate receiver for the common message and at the same time a non-legitimate receiver for the confidential message, we have to assume the worst channel realization to guarantee reliability of the common message and the best channel realization to guarantee secrecy of the confidential message.

Finally, we want to mention that the results derived for the compound BCC immediately yield also solutions for the corresponding multicast scenario. Here, the transmitter sends a common message to a whole group of receivers. Some of them are further legitimate receivers of a confidential message, which has to be kept secret from the other group of non-legitimate receivers. Then the compound BCC provides a framework which also includes this scenario. In this case, the number of possible channel realizations corresponds to the groups of legitimate and non-legitimate receivers so that each particular channel realization belongs to a certain legitimate or non-legitimate receiver.

Acknowledgment

The authors would like to thank for the valuable discussions with the Federal Office for Information Security (BSI) of Germany which raised the problem of robust broadcasting of common and confidential messages over compound channels.

References

[1] R. F. Wyrembelski and H. Boche, "Strong secrecy in compound broadcast channels with confidential messages," in *Proc. IEEE Int. Symp. Inf. Theory*, Cambridge, MA, USA, Jul. 2012, pp. 76–80.

[2] R. F. Schaefer and H. Boche, "Strong secrecy and decoding performance analysis for robust broadcasting under channel uncertainty," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Florence, Italy, May 2014, pp. 3973–3977.

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

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

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

[6] Y. Liang, H. V. Poor, and S. Shamai (Shitz), "Information theoretic security," *Found. Trends Commun. Inf. Theory*, vol. 5, nos. 4–5, pp. 355–580, 2009.

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

[8] E. A. Jorswieck, A. Wolf, and S. Gerbracht, "Secrecy on the physical layer in wireless networks," in *Trends in Telecommunications Technologies*. Rijeka, Croatia: InTech, Mar. 2010, pp. 413–435.

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

[10] 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.

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

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

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

[14] U. Maurer and S. Wolf, "Information-theoretic key agreement: From weak to strong secrecy for free," in *Advances in Cryptology—EUROCRYPT* (Lecture Notes in Computer Science), vol. 1807. Berlin, Germany: Springer-Verlag, May 2000, pp. 351–368.

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

[16] D. Blackwell, L. Breiman, and A. J. Thomasian, "The capacity of a class of channels," *Ann. Math. Statist.*, vol. 30, no. 4, pp. 1229–1241, Dec. 1959.

[17] J. Wolfowitz, "Simultaneous channels," *Archive Rational Mech. Anal.*, vol. 4, no. 1, pp. 371–386, 1960.

[18] Y. Liang, G. Kramer, H. V. Poor, and S. Shamai (Shitz), "Compound wiretap channels," *EURASIP J. Wireless Commun. Netw.*, vol. 2009, pp. 1–13, Mar. 2009, Art. ID 142374.

[19] E. Ekrem and S. Ulukus, "On Gaussian MIMO compound wiretap channels," in *Proc. 44th Annu. Conf. Inf. Sci. Syst.*, Princeton, NJ, USA, Mar. 2010, pp. 1–6.

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

[21] 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./Jul. 2009, pp. 1283–1287.

[22] P. Brémoud, *Markov Chains: Gibbs Fields, Monte Carlo Simulation, and Queues*. New York, NY, USA: Springer-Verlag, 1999.

[23] H. Boche and R. F. Schaefer, "Wiretap channels with side information—Strong secrecy capacity and optimal transceiver design," *IEEE Trans. Inf. Forensics Security*, vol. 8, no. 8, pp. 1397–1408, Aug. 2013.

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

[25] E. C. Song, E. Soljanin, P. Cuff, H. V. Poor, and K. Guan, "Rate-distortion-based physical layer secrecy with applications to multimode fiber," *IEEE Trans. Commun.*, vol. 62, no. 3, pp. 1080–1090, Mar. 2014.

[26] J. Körner and K. Marton, "General broadcast channels with degraded message sets," *IEEE Trans. Inf. Theory*, vol. 23, no. 1, pp. 60–64, Jan. 1977.

[27] R. F. Wyrembelski, I. Bjelaković, T. J. Oechtering, and H. Boche, "Optimal coding strategies for bidirectional broadcast channels under channel uncertainty," *IEEE Trans. Commun.*, vol. 58, no. 10, pp. 2984–2994, Oct. 2010.

[28] D. Dubhasi and A. Panconesi, *Concentration of Measure for the Analysis of Randomized Algorithms*. Cambridge, U.K.: Cambridge Univ. Press, 2009.

[29] R. Ahlswede and A. Winter, "Strong converse for identification via quantum channels," *IEEE Trans. Inf. Theory*, vol. 48, no. 3, pp. 569–579, Mar. 2002.

[30] P. C. Shields, *The Ergodic Theory of Discrete Sample Paths*. Providence, RI, USA: AMS, 1996.

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

Rafael F. Schaefer (S'08–M'12) 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 Post-Doctoral Research Fellow with the Department of Electrical Engineering, Princeton University, Princeton, NJ, USA. He 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.

Holger Boche (M'04–SM'07–F'11) received the Dipl.-Ing. and Dr.-Ing. degrees in electrical engineering from the Technische Universität Dresden, Dresden, Germany, in 1990 and 1994, respectively, the degree 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. He received the Dr.rer.nat. degree in pure mathematics from the Technische Universität Berlin, Berlin, Germany, in 1998. In 1997, he joined the Fraunhofer-Institut für Nachrichtentechnik, Heinrich-Hertz-Institut (HHI), Berlin. In 2002, he was a Full Professor of Mobile Communication Networks with the Institute for Communications Systems, Technische Universität Berlin. In 2003, he became the Director of the Fraunhofer German-Sino Laboratory for Mobile Communications, Berlin, Germany, and in 2004, he became the Director of HHI. Since 2010, he has been with the Institute of Theoretical Information Technology, and a Full Professor with the Technische Universität München, Munich, Germany. Since 2014, he has been a member and an Honorary Fellow of the TUM Institute for Advanced Study, Munich. He was a Visiting Professor with ETH Zurich, Zurich, Switzerland, he was a visiting professor during the winter terms 2004 and 2006, and the Royal Institute of Technology, Stockholm, Sweden, in Summer 2005. He is a member of the IEEE Signal Processing Society's Signal Processing and Communications and Signal Processing Theory and Methods technical committees. He was elected as a member of the German Academy of Sciences Leopoldina, in 2008, and the Berlin Brandenburg Academy of Sciences and Humanities in 2009. He was a recipient of the Technische Kommunikation Research Award from the Alcatel SEL Foundation in 2003, the Innovation Award from the Vodafone Foundation in 2006, and the Gottfried Wilhelm Leibniz Prize from the Deutsche Forschungsgemeinschaft (German Research Foundation) in 2008. He was a co-recipient of the 2006 IEEE Signal Processing Society Best Paper Award, and a recipient of the 2007 IEEE Signal Processing Society Best Paper Award.