Secure Broadcasting Using Independent Secret Keys

Rafael F. Schaefer , *Senior Member, IEEE*, Ashish Khisti, *Member, IEEE*, and H. Vincent Poor, *Fellow, IEEE*

(Invited Paper)

Abstract—The problem of secure broadcasting with independent secret keys is studied. The particular scenario is analyzed in which a common message has to be broadcast to two legitimate receivers, while keeping an external eavesdropper ignorant of it. The transmitter shares independent secret keys of sufficiently high rates with both legitimate receivers, which can be used in different ways: they can be used as one-time pads to encrypt the common message, as fictitious messages for wiretap coding, or as a hybrid of these. In this paper, capacity results are established when the broadcast channels involving the three receivers are degraded. If both legitimate channels are degraded versions of the eavesdropper's channel, it is shown that the one-time pad approach is optimal for several cases, yielding corresponding capacity expressions. Alternatively, the wiretap coding approach is shown to be optimal if the eavesdropper's channel is degraded with respect to both legitimate channels, establishing capacity in this case as well. If the eavesdropper's channel is neither the strongest nor the weakest, an intricate scheme that carefully combines both concepts of one-time pad and wiretap coding with fictitious messages turns out to be capacity-achieving. Finally we also obtain some results for the general non-degraded broadcast channel.

Index Terms—Broadcast channel, common message, secret key, secrecy capacity, one-time pad, wiretap coding, fictitious messages.

I. INTRODUCTION

SHANNON was the first who studied in [3] the problem of secure communication from an information theoretic perspective. He considered a noiseless communication scenario, in which transmitter and receiver share a secret key that is unknown to an eavesdropper. Used as a *one-time pad*, this secret key enables secure transmission of the confidential message.

Subsequently, Wyner considered the noisy case in [4], where he introduced the so-called *wiretap channel*. He extended

the problem studied by Shannon insofar that the legitimate receiver and the eavesdropper now observe noisy versions of the channel input. In addition, there is no secret key available as in [3] so that the communication must be secured solely by exploiting the properties of the noisy channel. Recently, this area of *information theoretic security* or *physical layer security* has drawn considerable attention especially in the area of wireless communication where it provides a promising complement to cryptographic approaches; see for example [5]–[9] and references therein. These concepts have been extended to several multi-user scenarios such as the broadcast channel [10]–[13], multiple access channel [14], [15], and interference channel [16]. All these works have in common that no secret key is available to the legitimate users.

These two approaches were combined in [17]–[19], which study the (noisy) wiretap channel with a shared secret key. This was done from a rate-distortion point of view in [17] and [18], while [19] established the secrecy capacity for the case in which no distortion is allowed at the legitimate receiver. The wiretap channel with a shared secret key and its main concepts and results are briefly reviewed in Section II. Related to this problem is the wiretap channel with secured feedback, as this feedback can be used to create a shared secret key [20]–[22].

A secret key shared between a transmitter and a receiver can be used to securely transmit to that receiver. If the rate of the secret key is sufficiently high, the problem becomes trivial in the single user setting (i.e., the wiretap channel with shared secret key) as the whole message can be encrypted with the secret key and the secrecy capacity is simply given by the regular channel capacity. However, when the transmitter shares independent secret keys with multiple receivers, a certain secret key can be used to securely transmit to the respective receiver, but might harm other receivers which do not share this key. Thus, multiple secret keys can result in conflicting payoffs at different receivers making it a challenging and non-trivial problem.

Surprisingly, to the best of our knowledge the use of multiple secret keys in noisy multi-user communication scenarios has not been studied previously. Accordingly, the question of secure communication in *broadcast channels (BC) with independent secret keys* is an interesting extension in this direction. In particular, this setting describes the problem of securely broadcasting a common message to two legitimate receivers, while keeping an eavesdropper ignorant of it. The transmitter shares independent secret keys of arbitrary key rates with the two receivers. This model is introduced in Section III.

Secure communication can now be realized by different approaches. As shared secret keys are available at the transmitter and both receivers, this suggests to follow [3] and use them as one-time pads to encrypt the common message. However, each receiver is aware of only one secret key. Thus, the more the secret key of one receiver is used to secure the message, the more the other receiver is hurt as the unknown secret key creates a useless message for it that interferes with the desired one. We show that when the eavesdropper has the "strongest" channel (i.e. both legitimate channels are degraded versions of it), the one-time pad approach is indeed capacity-achieving. More specifically, both secret keys are used to create two encrypted messages which are then accordingly encoded and transmitted with the help of a regular capacity-achieving (non-secure) broadcast channel code.

On the other hand, the properties of the noisy channels can be exploited by applying information theoretic secrecy concepts such as *wiretap coding* [5]–[9]. This approach is based on the idea of allocating some of the available rate for additional randomization to "confuse" the eavesdropper. The drawback of this is that it reduces the remaining rate available for the actual transmission of the confidential message. However, this approach turns out to be optimal if the eavesdropper's channel is the "weakest" among all channels (i.e. it is degraded with respect to both legitimate channels). To this end, the secret keys are used as *fictitious messages* for the randomization. As each receiver knows one of the secret keys, the randomization overhead to be decoded at the receivers is reduced.

While in the previous cases the eavesdropper has either the strongest or the weakest channel, the much more challenging and complicated scenario occurs when one legitimate channel is stronger and one is weaker than the eavesdropper's channel. In this case, it is shown that, to achieve capacity, a careful combination of one-time pad and wiretap coding with fictitious messages is needed. The converse involves a non-trivial "sum-rate"-like upper bound on the common message rate.

The above results are obtained in the context of discrete memoryless channels (DMCs). We further consider Gaussian channels, in which the legitimate receivers' and the eavesdropper's channel observations are impaired by additive Gaussian noise. The previous results are extended to this case resulting in a complete characterization of the capacity of the Gaussian BC with independent secret keys.

All results are stated in Section IV, while the corresponding proofs are delegated to Section V. A conclusion is then 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; $\mathbb{N}$ and $\mathbb{R}_+$ are the sets of positive integers and non-negative real numbers; $X - Y - Z$ denotes a Markov chain of random variables X, Y, and Z in this order; for a sequence $X^n = (X_1, X_2, \ldots, X_n)$ we frequently write $X^{i-1} = (X_1, X_2, \ldots, X_{i-1})$ and $X_{i+1}^n = (X_{i+1}, X_{i+2}, \ldots, X_n)$; all logarithms, exponentials, and information quantities are with
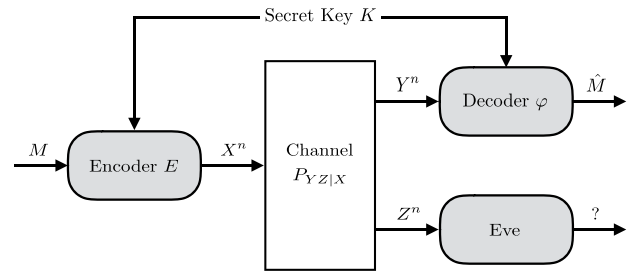

Fig. 1. Wiretap channel with a shared secret key.

respect to the base 2; $H(\cdot)$ and $I(\cdot; \cdot)$ are the traditional entropy and mutual information; $\mathcal{P}(\cdot)$ is the set of all probability distributions; $\oplus$ denotes modulo addition and $[a]^+$ abbreviates $\max\{a, 0\}$; lhs $:=$ rhs assigns the value of the right hand side (rhs) of an equation to its left hand side (lhs).

II. WIRETAP CHANNEL WITH SHARED SECRET KEY

In this section we review the *wiretap channel with shared secret key* as depicted in Fig. 1. This model generalizes and combines two important communication models: Shannon's cypher system [3] and the classical (noisy) wiretap channel as introduced by Wyner in [4]. Both lead to fundamentally different approaches of how secrecy is realized. These are presented and discussed as well.

A. System Model

Let $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ be finite input and output sets. For input and output sequences $x^n \in \mathcal{X}$, $y^n \in \mathcal{Y}^n$, and $z^n \in \mathcal{Z}^n$ of length n, the discrete memoryless wiretap channel is given by the transition probability $P_{YZ|X}^n(y^n, z^n|x^n) := \prod_{i=1}^n P_{YZ|X}(y_i, z_i|x_i)$.

The transmitter wishes to send a message M to the legitimate receiver while keeping the eavesdropper ignorant of it. The transmitter and legitimate receiver share a common secret key K of rate R_K which is unknown to the eavesdropper. The message M and the secret key K are assumed to be independent of each other and uniformly distributed over the sets $\mathcal{M} := \{1, \ldots, M_n\}$ and $\mathcal{K} := \{1, \ldots, K_n\}$.

Definition 1: An (n, M_n, K_n)-*code* for the wiretap channel with shared secret key consists of a (stochastic) encoder

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

and a decoder at the legitimate receiver

$$\varphi : \mathcal{Y}^n \times \mathcal{K} \to \mathcal{M}. \qquad (2)$$

Then the average probability of decoding error at the legitimate receiver is given by

$$\bar{e}_n = \frac{1}{|\mathcal{M}||\mathcal{K}|}$$
$$\times \sum_{m \in \mathcal{M}} \sum_{k \in \mathcal{K}} \sum_{x^n \in \mathcal{X}^n} \sum_{y^n : \varphi(y^n, k) \neq m} P_{Y|X}^n(y^n|x^n) E(x^n|m, k)$$
$$(3)$$

where $E(x^n|m, k)$ denotes the probability that the message $m \in \mathcal{M}$ and the secret key $k \in \mathcal{K}$ are encoded into the codeword $x^n \in \mathcal{X}^n$.

To ensure the confidentiality of the message, we require

$$I(M; Z^n) \leq \delta_n \tag{4}$$

for $\delta_n > 0$ with M the random variable uniformly distributed over the set of messages $\mathcal{M}$ and $Z^n = (Z_1, \ldots, Z_n)$ the channel output at the eavesdropper. This condition is termed *strong secrecy* [23], [24] and the motivation is to have the total amount of information leaked to the eavesdropper small.

Definition 2: For a given key rate $R_K > 0$, a rate $R > 0$ is an *achievable secrecy rate* for the wiretap channel with shared secret key if for every $\tau > 0$ there exists an $n(\tau) \in \mathbb{N}$ and a sequence of (n, M_n, K_n)-codes such that for all $n \geq n(\tau)$ we have $\frac{1}{n} \log M_n \geq R - \tau$, $\frac{1}{n} \log K_n = R_K$, and $I(M; Z^n) \leq \delta_n$ while $\bar{e}_n, \delta_n \to 0$ as $n \to \infty$. The *secrecy capacity C* is given by the supremum of all achievable secrecy rates R.

The wiretap channel with shared secret key has been studied in [17]–[19]. The secrecy capacity has been established and is restated in the following theorem.

Theorem 1 [19]: The secrecy capacity C of the wiretap channel with shared secret key of rate R_K is

$$C = \max_{P_{UVX}} \min \left\{ [I(V; Y|U) - I(V; Z|U)]^+ + R_K, I(V; Y) \right\} \tag{5}$$

such that $U - V - X - (Y, Z)$ form a Markov chain.

Depending on whether the term $I(V; Y|U) - I(V; Z|U)$ in (5) is positive or not, capacity can be achieved by using two fundamentally different approaches for realizing secrecy: the one-time pad approach or the principle of wiretap coding with fictitious messages. Both ingredients become most evident by considering either reversely degraded channels, i.e., $X - Z - Y$ form a Markov chain, or degraded channels, i.e., $X - Y - Z$ form a Markov chain. In these cases, the secrecy capacity simplifies as given in the following.

Corollary 1: The secrecy capacity C of the reversely degraded wiretap channel with shared secret key of rate R_K is

$$C = \max_{P_X} \min \left\{ R_K, I(X; Y) \right\} \tag{6}$$

if $X - Z - Y$ form a Markov chain.

Corollary 2: The secrecy capacity C of the degraded wiretap channel with shared secret key of rate R_K is

$$C = \max_{P_X} \min \left\{ I(X; Y) - I(X; Z) + R_K, I(X; Y) \right\} \tag{7}$$

if $X - Y - Z$ form a Markov chain.

Next we discuss both approaches of one-time pad and wiretap coding in detail and show how they achieve capacity in the respective scenarios.

B. One-Time Pad

In this approach, we use the shared secret key for encryption to secure the message transmission. Following the thinking of [3], the secret key can be used as a one-time pad, for which the secret key and message must be of the same rate, so that the secrecy capacity is immediately limited by the secret key

rate R_K. Then, the encrypted message is obtained based on a modulo-$|\mathcal{K}|$ addition as

$$M^\oplus = M \oplus K. \tag{8}$$

The encrypted message is then encoded using a classical channel encoder and transmitted over the noisy channel so that the secrecy capacity is further limited by the regular channel capacity, i.e., $\max_{P_X} I(X; Y)$. Having decoded the encrypted message $\hat{M}^\oplus$, the legitimate receiver can then use the secret key to recover the original message $\hat{M}$, i.e., $\hat{M}^\oplus \oplus K = \hat{M} \oplus K \oplus K = \hat{M}$. Since M and K are uniformly distributed and independent of each other, M is perfectly secret from the eavesdropper, i.e., $I(M; Z^n) = 0$, even if it would be able to decode the encrypted message $M^\oplus$. We conclude that the one-time pad approach achieves the rate

$$R = \max_{P_X} \min \left\{ R_K, I(X; Y) \right\}. \tag{9}$$

We observe that the rate in (9) is independent of the eavesdropper's channel and depend only on the available key rate and the legitimate channel. Accordingly, this approach works for any configuration of legitimate and eavesdropper's channels.

In particular, in the case of reversely degraded channels, i.e., $X - Z - Y$, the one-time pad approach is optimal as it achieves the secrecy capacity as given in Corollary 1. However, we observe that in the case of degraded channels, i.e., $X - Y - Z$, the one-time pad approach is suboptimal as the achieved rate (9) is in general strictly smaller than the secrecy capacity as given in Corollary 2 since $I(X; Y) - I(X; Z) > 0$.

Remark 1: We immediately also obtain the corresponding result for Shannon's cypher system, which corresponds to the noiseless case, i.e., $X = Y = Z$. The secrecy capacity C of the noiseless wiretap channel with shared secret key of rate R_K is

$$C = \max_{P_X} \min \left\{ R_K, H(X) \right\}. \tag{10}$$

Remark 2: For sufficiently large key rate, we see from (6) and (10) that the secrecy capacity is limited by the regular channel capacity, i.e., $C = \max_{P_X} I(X; Y)$ and $C = \max_{P_X} H(X)$ respectively. The influence of the eavesdropper on the capacity completely vanishes in this case.

C. Wiretap Coding With Fictitious Messages

In contrast to the one-time pad approach, this approach exploits the noisy channel to realize secure communication. If there is no secrecy requirement, all available resources are used for message transmission. To incorporate secrecy, the crucial idea of wiretap coding [5]–[9] is now to use a certain part for randomization by introducing "dummy" messages. This induces multiple valid codewords for each confidential message and the one used for the actual transmission is chosen uniformly at random. The key insight is now to choose the randomization rate (i.e. the rate of the dummy message) for each confidential message according to the eavesdropper's channel quality, i.e., roughly as $I(X; Z)$. This will saturate the eavesdropper with useless information carried by the dummy

messages leaving no remaining resources for decoding the confidential message itself [25].

In the following we present a wiretap coding scheme and show how the secret key is incorporated. Let us assume that $I(X; Y) - I(X; Z) > 0$ and $R_K \leq I(X; Z)$. For any input distribution $P_X \in \mathcal{P}(X)$ we generate $|\mathcal{M}||\mathcal{K}||\mathcal{J}|$ independent codewords $x_{mkj}^n \in \mathcal{X}^n$ where

$$|\mathcal{M}| < 2^{n(I(X;Y)-I(X;Z)+R_K-2\epsilon)} \tag{11a}$$
$$|\mathcal{K}| = 2^{nR_K} \tag{11b}$$
$$|\mathcal{J}| > 2^{n(I(X;Z)-R_K+\epsilon)}. \tag{11c}$$

The crucial idea here is to treat the secret key as a fictitious message for additional randomization. As the size of the secret key and the dummy randomization messages satisfy

$$\frac{1}{n}\log(|\mathcal{K}||\mathcal{J}|) > I(X; Z) + \epsilon, \tag{12}$$

we have enough resources allocated to saturate the eavesdropper to ensure that $I(M; Z^n) \leq \delta_n$ holds, i.e., strong secrecy (4) is satisfied, cf. for example [26]–[28].

It remains to verify that the legitimate receiver can reliably decode the confidential message. Since the legitimate receiver has the secret key $k \in \mathcal{K}$ as side information available, the remaining unknown indices are $m \in \mathcal{M}$ and $j \in \mathcal{J}$ which are appropriately chosen to satisfy

$$|\mathcal{M}||\mathcal{J}| \leq 2^{n(I(X;Y)-\epsilon)}. \tag{13}$$

It is easy to verify that this enables the legitimate receiver to reliably decode the remaining indices.

In particular, in the case of degraded channels, i.e., $X - Y - Z$ so that $I(X; Y) - I(X; Z) > 0$, wiretap coding with fictitious messages is optimal as it achieves the secrecy capacity as given in Corollary 2. However, we observe that in the case of reversely degraded channels, i.e., $X - Z - Y$, wiretap coding with fictitious messages is in general suboptimal (except for sufficiently large key rates for which the secrecy capacity becomes $\max_{P_X} I(X; Y)$).

Remark 3: From (12) we observe that the maximum useful secret key rate is $I(X; Z)$. Indeed, even if $R_K > I(X; Z)$, it suffices to only use $R_K = I(X; Z)+\epsilon$ as this already saturates the eavesdropper's channel. The secrecy capacity in this case is then $\max_{P_X} I(X; Y)$ and there is no need for additional dummy messages for randomization in this case, i.e., $\mathcal{J} = \emptyset$.

Remark 4: If $I(X; Y) - I(X; Z) < 0$, e.g., for reversely degraded channels $X - Z - Y$, wiretap coding with fictitious messages in general performs worse than the one-time pad approach. In particular, if $I(X; Z) \geq I(X; Y)+R_K$, no secure communication is possible at all, i.e., $R = 0$, while the one-time pad approach still achieves the rate $R = \max_{P_X} I(X; Y)$ in this case.

From the previous reasoning and particularly Theorem 1 and Corollary 2 we obtain the corresponding secrecy capacity for the classical wiretap channel without shared key, i.e., $R_K = 0$, as stated next.

Corollary 3: The secrecy capacity C of the wiretap channel is

$$C = \max_{P_{VX}} \left[I(V; Y) - I(V; Z) \right] \tag{14}$$

such that $V - X - (Y, Z)$ form a Markov chain. If the wiretap channel is degraded, i.e., $X - Y - Z$ form a Markov chain, the secrecy capacity C becomes

$$C = \max_{P_X} \left[I(X; Y) - I(X; Z) \right]. \tag{15}$$

D. A Combination of Both

The previous discussion has shown that neither of the two approaches of one-time pad and wiretap coding with fictitious messages is superior to the other. In fact, whenever the eavesdropper's channel is "stronger" than the legitimate channel $(I(X; Y) - I(X; Z) < 0)$ the one-time pad performs better than the wiretap coding approach and vice versa when the eavesdropper's channel is weaker $(I(X; Y) - I(X; Z) > 0)$.

In the following, we present a scheme that combines both approaches of one-time pad and wiretap coding. To make things non-trivial, we assume $R_K < I(X; Y)$, since otherwise the secrecy capacity is $C = \max_{P_X} I(X; Y)$ as discussed above. The main idea is now to use the secret key as a one-time pad to encrypt as much of the message as possible and use the wiretap coding approach to protect the remaining part of the message. Accordingly, we split the message M into two parts

$$M = (M', M^{\oplus}) \tag{16}$$

with rates R' and R_K, where the part M' will be protected by the wiretap coding approach and $M^{\oplus}$ by the one-time pad approach. For any input distribution $P_X \in \mathcal{P}(X)$ we generate $|\mathcal{M}'||\mathcal{M}^{\oplus}||\mathcal{J}|$ independent codewords $x_{m'm^{\oplus}j}^n \in \mathcal{X}^n$ where

$$|\mathcal{M}'| < 2^{n([I(X;Y)-\max\{I(X;Z),R_K\}]^+ -2\epsilon)} \tag{17a}$$
$$|\mathcal{M}^{\oplus}| = 2^{nR_K} \tag{17b}$$
$$|\mathcal{J}| = \begin{cases} 2^{n([I(X;Z)-R_K]^+ +\epsilon)} & \text{if } I(X; Y) > I(X; Z) \\ 0 & \text{if } I(X; Y) \leq I(X; Z). \end{cases} \tag{17c}$$

First, we observe that whenever $I(X; Y) - I(X; Z) \leq 0$, the sets $\mathcal{M}'$ and $\mathcal{J}$ are empty and we are falling back to the one-time pad approach as discussed in Section II-B. If $I(X; Y) - I(X; Z) > 0$, we use both approaches and secrecy is ensured similarly as in (12) and (13). To be specific, the sets are chosen appropriately to satisfy

$$\frac{1}{n}\log(|\mathcal{M}^{\oplus}||\mathcal{J}|) > I(X; Z) + \epsilon, \tag{18}$$

which gives us secrecy, and further

$$|\mathcal{M}'||\mathcal{M}^{\oplus}||\mathcal{J}| \leq 2^{n(I(X;Y)-\epsilon)}, \tag{19}$$

which gives us reliability.

Remark 5: For both wiretap coding with fictitious messages, cf. Section II-C, and the combined approach, the secret key contributes to the equivocation of the eavesdropper by saturating its channel. But note that in wiretap coding with fictitious messages the secret key does not carry any useful information while in the combined approach it further carries part of the confidential message.

Remark 6: For the case $I(X; Y) - I(X; Z) > 0$, both approaches have the same performance in terms of achievable secrecy rate. However, they fundamentally differ in
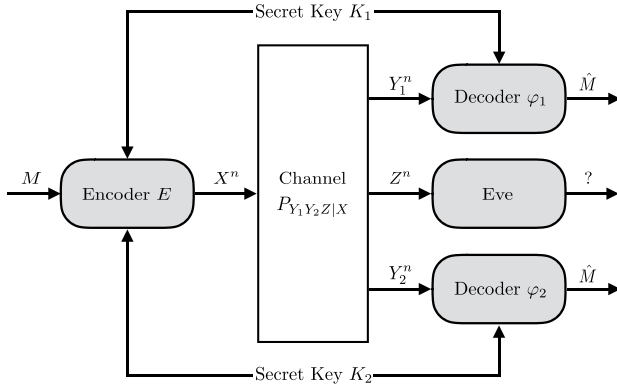
Fig. 2. Broadcast channel in which the transmitter shares independent secret keys K_1 and K_2 with legitimate receivers 1 and 2 respectively. A common message M is sent to both legitimate receivers keeping the eavesdropper ignorant of it.

the construction and their codebooks are of different size. In particular, while the overall codebook size of the combined approach is matched to the legitimate channel (as the receiver has to decode all three indices), wiretap coding with fictitious messages uses a codebook whose size is larger than the legitimate channel would allow, but the receiver can use its secret key as side information to reduce it to a suitable subset.

III. BC WITH INDEPENDENT SECRET KEYS

In the remainder of this paper we study the *broadcast channel (BC) with independent secret keys* as depicted in Fig. 2. Let $\mathcal{X}$, $\mathcal{Y}_1$, $\mathcal{Y}_2$, and $\mathcal{Z}$ be finite input and output sets. For input and output sequences $x^n \in \mathcal{X}^n$, $y_1^n \in \mathcal{Y}_1^n$, $y_2^n \in \mathcal{Y}_2^n$, and $z^n \in \mathcal{Z}^n$ of length n, the discrete memoryless broadcast channel is given by the transition probability $P_{Y_1 Y_2 Z|X}^n(y_1^n, y_2^n, z^n|x^n) \coloneqq \prod_{i=1}^n P_{Y_1 Y_2 Z|X}(y_{1,i}, y_{2,i}, z_i|x_i)$.

Throughout the rest of the paper we study the particular communication scenario in which the transmitter broadcasts a common message M to receivers 1 and 2, while keeping the eavesdropper ignorant of it. The transmitter shares independent secret keys K_1 and K_2 with receivers 1 and 2. The message and both keys are assumed to be independent of each other and uniformly distributed over the sets $\mathcal{M} \coloneqq \{1, \ldots, M_n\}$ and $\mathcal{K}_i \coloneqq \{1, \ldots, K_{i,n}\}$, $i = 1, 2$. We also write $K_{12} = (K_1, K_2)$ and $\mathcal{K}_{12} = \mathcal{K}_1 \times \mathcal{K}_2$ for short.

Definition 3: An $(n, M_n, K_{1,n}, K_{2,n})$-*code* for the BC with independent secret keys consists of a (stochastic) encoder

$$E : \mathcal{M} \times \mathcal{K}_1 \times \mathcal{K}_2 \to \mathcal{P}(\mathcal{X}^n) \qquad (20)$$

and decoders at receivers 1 and 2

$$\varphi_1 : \mathcal{Y}_1^n \times \mathcal{K}_1 \to \mathcal{M}, \qquad (21a)$$
$$\varphi_2 : \mathcal{Y}_2^n \times \mathcal{K}_2 \to \mathcal{M}. \qquad (21b)$$

Then the average probability of decoding error at receiver i, $i = 1, 2$, is given by

$$\bar{e}_{i,n} = \frac{1}{|\mathcal{M}||\mathcal{K}_{12}|}$$
$$\times \sum_{m \in \mathcal{M}} \sum_{k_{12} \in \mathcal{K}_{12}} \sum_{x^n \in \mathcal{X}^n} \sum_{y_i^n : \varphi_i(y_i^n, k_i) \neq m} P_{Y_i|X}^n(y_i^n|x^n)$$
$$\times E(x^n|m, k_{12}). \qquad (22)$$

As for the wiretap channel with a shared secret key, we require strong secrecy of the common message, i.e.,

$$I(M; Z^n) \leq \delta_n \qquad (23)$$

with $\delta_n > 0$.

Definition 4: A rate $R > 0$ is an *achievable secrecy rate* for the BC with independent secret keys if for every $\tau > 0$ there exists an $n(\tau) \in \mathbb{N}$, key rates $R_1 > 0$, $R_2 > 0$, and a sequence of $(n, M_n, K_{1,n}, K_{2,n})$-codes such that for all $n \geq n(\tau)$ we have $\frac{1}{n} \log M_n \geq R - \tau$, $\frac{1}{n} \log K_{i,n} = R_i$, $i = 1, 2$, and $I(M; Z^n) \leq \delta_n$ while $\bar{e}_{1,n}, \bar{e}_{2,n}, \delta_n \to 0$ as $n \to \infty$. The *secrecy capacity* C is given by the supremum of all achievable secrecy rates R.

Remark 7: Note that in contrast to Definition 2, the rates of the secret keys are not fixed and specified a priori. In particular, Definition 4 allows us to choose arbitrary key rates R_1 and R_2. In other words, the key rates are sufficiently high in the sense that we are provided with as much secret key as we desire.

From the problem definition, in principle there are different methods possible to keep the common message secret. The shared secret keys suggest that we might use a one-time pad approach which protects the message with the help of the secret keys similarly as discussed in Section II-B. On the other hand, the transmitter can exploit the nature of the wireless channel by using a channel code based on the idea of wiretap coding as discussed in Section II-C. In the following we will explore these different approaches and show that which particular approach should be used depends on the channel conditions.

IV. MAIN RESULTS

In this section we present the main results for the BC with independent secret keys and discuss the insights that we obtain from them. The corresponding proofs are relegated to Section V and the appendix, respectively.

A. Equal Channel Outputs

We start with the simplest scenario, in which both legitimate receivers and the eavesdropper receive signals of the same quality and, in particular, $Y = Y_1 = Y_2 = Z$.

If a legitimate receiver has no advantage over the eavesdropper at the physical layer, wiretap coding as discussed in Section II-C cannot be used. Accordingly, to secure the message transmission, we use the secret keys as one-time pads as presented in Section II-B to create two *encrypted* messages

$$M_1 = M \oplus K_1, \quad \text{and} \quad M_2 = M \oplus K_2. \qquad (24)$$

These two messages are independent of each other and then encoded using a classical broadcast channel encoder and transmitted to the corresponding receivers. Having decoded the encrypted messages M_1 and M_2, each receiver can then use its own secret key to obtain the desired original message $\hat{M}$, i.e., $\hat{M}_1 \oplus K_1 = \hat{M} \oplus K_1 \oplus K_1 = \hat{M}$ and $\hat{M}_2 \oplus K_2 = \hat{M} \oplus K_2 \oplus K_2 = \hat{M}$ respectively. It becomes apparent that this one-time pad approach requires a total secret key rate of twice
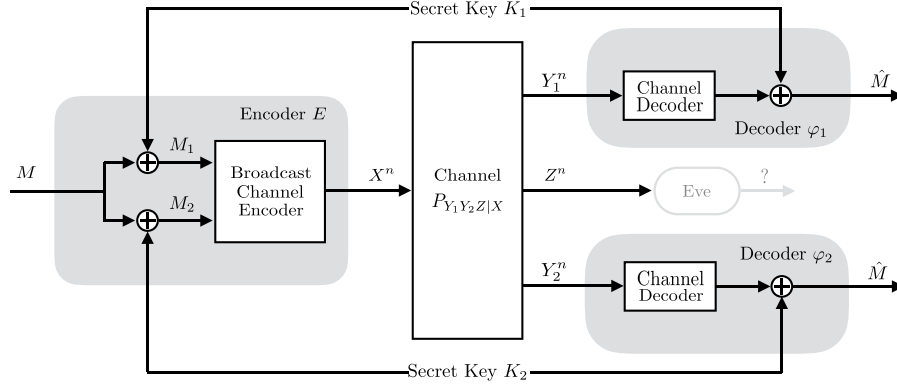
Fig. 3. One-time pad approach to the BC with independent secret keys. The shared keys K_1 and K_2 are used to encrypt the common message M by creating two independent messages M_1 and M_2. These two messages are encoded by a classical broadcast channel encoder and transmitted over the channel. Having decoded the individual messages $\hat{M}_1$ and $\hat{M}_2$, the legitimate receivers obtain the desired common message $\hat{M}$ with the help of the shared keys.

the common message rate. Since the message M and both keys K_1 and K_2 are uniformly distributed and independent of each other, M is kept perfectly secret from the eavesdropper, i.e., $I(M; Z^n) = 0$, even if it would be able to decode the encrypted messages M_1 or M_2. This encoding-decoding process is depicted in Fig. 3.

For the case of equal channel outputs we then obtain the following result.

Theorem 2: The secrecy capacity C of the BC with independent secret keys for equal channel outputs is

$$C = \max_{P_X} \frac{1}{2} I(X; Y), \qquad (25)$$

with $Y = Y_1 = Y_2 = Z$, *i.e., time-sharing between both legitimate receivers is optimal.*

Proof: Achievability follows from using the secret keys as one-time pads to create encrypted individual messages as in (24) and time-sharing between these two encrypted messages. The converse follows from Fano's inequality and the standard properties of entropy and mutual information to exploit the tension from using independent secret keys from the two receivers. A detailed proof can be found in Section V-A. ■

From this we immediately obtain the result for the noiseless case where the output at all receivers equals the input, i.e., $X = Y_1 = Y_2 = Z$.

Corollary 4: The secrecy capacity C of the noiseless BC with independent secret keys is

$$C = \max_{P_X} \frac{1}{2} H(X), \qquad (26)$$

with $X = Y_1 = Y_2 = Z$, *i.e., time-sharing between both legitimate receivers is optimal.* ■

Remark 8: These results show that in the case of equal channel outputs the simple strategy of time-sharing is already capacity-achieving. Thereby, the secret keys are used as one-time pads to transform the common message into multiple individual encrypted messages.

B. Strongest Eavesdropper

After having obtained some first insights for the case of equal channel outputs, we now turn to the case in which all channel outputs are of different quality but the eavesdropper has the strongest channel among all receivers in a certain sense.

Let us first assume that the channels are reversely degraded, i.e., we have the following Markov chain: $X - Z - Y_1 - Y_2$.

Theorem 3: The secrecy capacity of the BC with independent secret keys for reversely degraded channels is

$$C = \max_{P_{UX}} \min \left\{ \begin{array}{c} I(X; Y_1|U) \\ I(U; Y_2) \end{array} \right\}, \qquad (27)$$

such that $U - X - Z - Y_1 - Y_2$ *form a Markov chain, i.e., superposition coding is optimal. Further, the cardinality of the range of U can be bounded by* $|\mathcal{U}| \leq |\mathcal{X}| + 1$.

Proof: Achievability follows from using the secret keys as one-time pads to create encrypted individual messages and then using superposition coding. The converse follows from Fano's inequality and standard arguments for the degraded broadcast channel. A detailed proof is given in Section V-B. ■

Remark 9: The fact that the eavesdropper's channel is the strongest among all channels (in the sense that both legitimate channels are degraded versions of it) suggests to use the secret keys as one-time pads to secure the message as classical wiretap coding approaches will not work. In addition, the fact that the legitimate channels themselves can be ordered due to their degradedness suggests then to use a superposition coding scheme (as for the classical degraded BC). The previous result shows that this reasoning indeed yields a capacity-achieving coding scheme.

We have seen that using the secret keys in this way essentially turns the problem of securely broadcasting a common message into the problem of broadcasting two independent individual messages. Having this observation in mind, we immediately obtain the following achievable rate which is based on the idea of Marton coding for the classical BC, cf. [29, Sec. 8.4].

Lemma 1: An achievable secrecy rate for the BC with independent secret keys is given by

$$R = \min \left\{ \begin{array}{l} I(U, W; Y_1), \\ I(V, W; Y_2), \\ \frac{1}{2}[I(U, W; Y_1) + I(V; Y_2|W) - I(U; V|W)], \\ \frac{1}{2}[I(V, W; Y_2) + I(U; Y_1|W) - I(U; V|W)] \end{array} \right\}$$
(28)

for any $P_{UVWX}(u, v, w, x)$ such that $(U, V, W) - X - (Y_1, Y_2, Z)$ form a Markov chain.

Proof: A sketch of the proof is provided in Appendix A. ∎

We want to stress that Lemma 1 yields an achievable secrecy rate for the BC with independent secret keys for the general case, i.e., *no* constraints on the channels to the legitimate receivers and the eavesdropper are imposed.

The previous Theorem 3 assumed that the channels are reversely degraded. Let us now generalize this discussion in the sense that we do not impose any ordering between the legitimate channels Y_1 and Y_2. But we still assume that the eavesdropper's channel is the "strongest" of all channels in the sense that the following Markov chains are satisfied: $X - Z - Y_1$ and $X - Z - Y_2$. Then we obtain the following outer bound which is somewhat similar to the UV-outer bound [30] or the Körner-Marton outer bound [31].

Theorem 4: The secrecy capacity C of the BC with independent secret keys for reversely degraded channels $X - Z - (Y_1, Y_2)$ is upper bounded by

$$C \leq \min \left\{ \begin{array}{l} I(U; Y_1), \\ I(V; Y_2), \\ \frac{1}{2}\big[I(U; Y_1) + I(X; Y_2|U)\big], \\ \frac{1}{2}\big[I(V; Y_2) + I(X; Y_1|V)\big] \end{array} \right\}$$
(29)

for all $P_{UVX}(u, v, x)$ such that $(U, V) - X - Z - (Y_1, Y_2)$ form a Markov chain. Further, the cardinalities of the ranges of U and V can be bounded by $|\mathcal{U}| \leq |\mathcal{X}| + 1$ and $|\mathcal{V}| \leq |\mathcal{X}| + 1$.

Proof: The proof of the upper bound in (29) resembles those of the UV-bound [30] and Körner-Marton bound [31]. A detailed proof is given in Section V-C. ∎

Remark 10: Similar to the achievability schemes discussed before, the UV-type upper bound indicates that in the case of a strongest eavesdropper, the communication problem resembles the structure of the BC with two independent individual messages (stemming from the two independent secret keys). Accordingly, similar to the classical UV-outer bound for the BC with two individual messages, our UV-type upper bound becomes tight for the previous special cases of equal channel outputs and reversely degraded channels. Thus, it recovers the converses of the corresponding Theorems 2 and 3.

Finally, we want to address the case of deterministic broadcast channels, for which the channel outputs at the legitimate receivers Y_1 and Y_2 are functions of the input X, i.e. $Y_1 = \phi_1(X)$ and $Y_2 = \phi_2(X)$. For reversely degraded channels $X - Z - (Y_1, Y_2)$ we then obtain the following.

Theorem 5: The secrecy capacity C of the BC with independent secret keys for reversely degraded channels $X - Z - (Y_1, Y_2)$, for which Y_1 and Y_2 are deterministic functions of X, is given by

$$C = \max_{P_X} \min \left\{ \begin{array}{l} H(Y_1), \\ H(Y_2), \\ \frac{1}{2}H(Y_1, Y_2) \end{array} \right\}.$$
(30)

Proof: Achievability follows from the Marton coding appraoch in Lemma 1 and the converse follows from the UV-type upper bound in Theorem 4. A detailed proof is given in Section V-D. ∎

C. Weakest Eavesdropper

Next, we study the opposite case in which the eavesdropper has the weakest channel among all receivers in a certain sense. As the legitimate receivers now have an advantage over the eavesdropper at the physical layer, it is possible to use wiretap coding principles as discussed in Section II-C. Recall that the basic idea of wiretap coding is not to use all available rate for transmitting the message, but to allocate some of the rate to "confuse" the eavesdropper by applying randomized encoding strategies [5]–[9]. If a sufficient amount of rate is spent for confusion, the eavesdropper will not be able to decode the transmitted message. Obviously, the more rate is allocated to this confusion, the less rate is available for the actual transmission of the message. Here is where the shared secret keys enter the picture in this approach. They will be treated as fictitious messages playing the role of the randomization resources which are (partly) available at the legitimate receivers. This approach is visualized in Fig. 4.

In the following we consider the case in which the eavesdropper has the "worst" channel in the sense that its channel is degraded with respect to both legitimate channels. Thus, we assume Markov chains $X - Y_1 - Z$ and $X - Y_2 - Z$; however, we impose no ordering between the legitimate channels themselves.

Theorem 6: The secrecy capacity C of the BC with independent secret keys and degraded channels $X - Y_1 - Z$ and $X - Y_2 - Z$ is

$$C = \max_{P_X} \min \left\{ \begin{array}{l} I(X; Y_1), \\ I(X; Y_2), \\ \frac{1}{2}[I(X; Y_1) + I(X; Y_2) - I(X; Z)] \end{array} \right\}.$$
(31)

Proof: Achievability follows by treating the secret keys as fictitious messages as part of the wiretap code. The converse follows by Fano's inequality applied to a suitably chosen sum-rate like bound. A detailed proof is given in Section V-E. ∎

Remark 11: An interesting observation is that in a capacity-achieving coding scheme, the sum rate of both secret keys must be equal to the eavesdropper's channel quality. This saturation leaves then no remaining resources for decoding of the actual common message.

Next, we want to weaken the "worst" eavesdropper channel assumption in the sense that we study the case of less noisy channels instead of degraded channels.
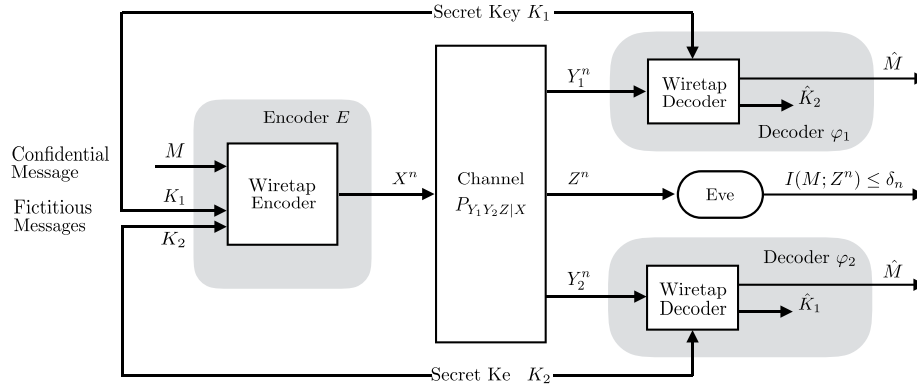
Fig. 4. Wiretap coding approach to the BC with independent secret keys. The shared keys K_1 and K_2 are treated as fictitious messages and are used by the wiretap encoder as randomization resources. The wiretap decoders each have one fictitious message as side information available so that it remains for them to decode the other (useless) fictitious message and the desired common message M.

Definition 5: A BC given by its joint probability $P_{YZ|X}$ is said to be *less noisy* if

$$I(U; Y) \geq I(U; Z) \tag{32}$$

holds for all $P_{UX}(u, x)$ such that $U - X - (Y, Z)$ form a Markov chain. In this case we say receiver Y is less noisy than receiver Z and write $Y \succeq Z$.

Accordingly, we assume $Y_1 \succeq Z$ and $Y_2 \succeq Z$ and, again, we impose no ordering between the legitimate channels themselves. The following shows that the capacity result given in Theorem 6 for degraded channels also holds for less noisy channels.

Theorem 7: The secrecy capacity C of the BC with independent secret keys and less noisy channels $Y_1 \succeq Z$ and $Y_2 \succeq Z$ is

$$C = \max_{P_X} \min \left\{ \begin{array}{l} I(X; Y_1), \\ I(X; Y_2), \\ \frac{1}{2}[I(X; Y_1) + I(X; Y_2) - I(X; Z)] \end{array} \right\}. \tag{33}$$

Proof: Achievability follows immediately from Theorem 6. The converse is an adaptation of that used in Theorem 6 incorporating the less noisy condition. A detailed proof is given in Section V-F. ∎

D. Eavesdropper in the Middle

The previous considerations either used the secret keys solely as one-time pads or solely as fictitious messages. In particular, the first approach is optimal if the eavesdropper has the "strongest" channel and the second approach is optimal if the eavesdropper has the "weakest" channel. The analysis becomes much more involved if the eavesdropper's channel is neither the strongest nor the weakest.

Accordingly, we study what happens if the eavesdropper's channel is in the middle (i.e. $X - Y_1 - Z - Y_2$ form a Markov chain). That is, it is stronger than one legitimate channel, but weaker than the other one. In this case, a combination of one-time pad and fictitious messages is capacity-achieving.

Theorem 8: The secrecy capacity C of the BC with independent secret keys for $X - Y_1 - Z - Y_2$ is

$$C = \max_{P_{UX}} \min \left\{ \begin{array}{l} I(U; Y_2), \\ \frac{1}{2}[I(X; Y_1) + I(U; Y_2) - I(U; Z)] \end{array} \right\} \tag{34}$$

such that $U - X - Y_1 - Z - Y_2$ form a Markov chain. The cardinality of the range of U can be bounded by $|\mathcal{U}| \leq |\mathcal{X}|+1$.

Proof: Achievability follows by applying a combination of the one-time pad approach and the fictitious messages approach. In particular, a superposition coding scheme is used in which the auxiliary random variable U serves as a "cloud center" protecting the confidential message using the fictitious messages approach. On top of that, "satellite codewords" for the stronger receiver are superimposed protected by the one-time pad approach. The converse follows from Fano's inequality applied to a suitably chosen sum-rate like bound with a careful identification of the auxiliary random variable. A detailed proof is given in Section V-G. ∎

E. Gaussian Channels

The results above can be extended to Gaussian channels, where input-output relations are given by

$$Y_1 = X + N_1, \tag{35}$$
$$Y_2 = X + N_2, \tag{36}$$
$$Z = X + N_Z, \tag{37}$$

where the additive Gaussian noise terms N_1, N_2, and N_Z have zero means and variances σ_1^2, σ_2^2, and σ_Z^2, i.e., $N_i \sim \mathcal{N}(0, \sigma_i^2)$, $i \in \{1, 2, Z\}$. We consider an average transmit power constraint P.

Since Gaussian broadcast channels are inherently degraded, the relation among the noise variances σ_1^2, σ_2^2, and σ_Z^2 immediately determines the degradedness order. Accordingly, we extend the previous results to completely characterize the Gaussian case. W.l.o.g. we assume $\sigma_1^2 \leq \sigma_2^2$, making receiver 1 the stronger of the two legitimate receivers.

1) Strongest Eavesdropper ($\sigma_Z^2 \leq \sigma_1^2 \leq \sigma_2^2$): The eavesdropper having the smallest noise power, i.e., the ordering is $\sigma_Z^2 \leq \sigma_1^2 \leq \sigma_2^2$, is analogous to the Markov chain condition $X - Z - Y_1 - Y_2$. Then the eavesdropper has the strongest channel as discussed in Section IV-B. In this case, the secrecy capacity is established in the following theorem.

Theorem 9: The secrecy capacity C of the Gaussian BC with independent secret keys for $\sigma_Z^2 \leq \sigma_1^2 \leq \sigma_2^2$ is

$$C = \max_{0 \leq \alpha \leq 1} \min \left\{ \begin{array}{l} \frac{1}{2} \log \left(1 + \frac{\alpha P}{\sigma_1^2}\right), \\ \frac{1}{2} \log \left(1 + \frac{(1-\alpha)P}{\alpha P + \sigma_2^2}\right) \end{array} \right\}. \quad (38)$$

Proof: Achievability follows immediately from Theorem 3 by extending the achievability scheme from the discrete memoryless case to the Gaussian case. The converse follows also from Theorem 3 and from the converse for the classical Gaussian BC, cf. for example [29, Th. 5.3]. The details are omitted for brevity. ∎

2) Weakest Eavesdropper ($\sigma_1^2 \leq \sigma_2^2 \leq \sigma_Z^2$): The eavesdropper having the highest noise power, i.e., the ordering is $\sigma_1^2 \leq \sigma_2^2 \leq \sigma_Z^2$, is analogous to the Markov chain condition $X - Y_1 - Y_2 - Z$. Then the eavesdropper has the weakest channel as discussed in Section IV-C. In this case, the secrecy capacity is established in the following theorem.

Theorem 10: The secrecy capacity C of the Gaussian BC with independent secret keys for $\sigma_1^2 \leq \sigma_2^2 \leq \sigma_Z^2$ is

$$C = \min \left\{ \begin{array}{l} \frac{1}{2} \log \left(1 + \frac{P}{\sigma_2^2}\right), \\ \frac{1}{4} \left[\log \left(1 + \frac{P}{\sigma_1^2}\right) + \log \left(1 + \frac{P}{\sigma_2^2}\right) \right. \\ \left. \qquad - \log \left(1 + \frac{P}{\sigma_Z^2}\right) \right] \end{array} \right\}. \quad (39)$$

Proof: Achievability follows immediately from Theorem 6 by extending the achievability scheme from the discrete memoryless case to the Gaussian case. The converse follows also from Theorem 6 together with the ideas for the converse of the classical Gaussian wiretap channel, cf. for example [29, Example 22.2]. The details are omitted for the sake of brevity. ∎

3) Eavesdropper in the Middle ($\sigma_1^2 \leq \sigma_Z^2 \leq \sigma_2^2$): When the eavesdropper's noise power satisfies $\sigma_1^2 \leq \sigma_Z^2 \leq \sigma_2^2$, it is analogous to the Markov chain condition $X - Y_1 - Z - Y_2$. Then the eavesdropper has neither the strongest nor the weakest channel as discussed in Section IV-D. In this case, the secrecy capacity is established in the following theorem.

Theorem 11: The secrecy capacity C of the Gaussian BC with independent secret keys for $\sigma_1^2 \leq \sigma_Z^2 \leq \sigma_2^2$ is

$$C \leq \max_{0 \leq \alpha \leq 1} \min \left\{ \begin{array}{l} \frac{1}{2} \log \left(1 + \frac{\alpha P}{(1-\alpha)P + \sigma_2^2}\right), \\ \frac{1}{4} \left[\log \left(1 + \frac{P}{\sigma_1^2}\right) + \log \left(1 + \frac{\alpha P}{(1-\alpha)P + \sigma_2^2}\right) \right. \\ \left. \qquad - \log \left(1 + \frac{\alpha P}{(1-\alpha)P + \sigma_Z^2}\right) \right] \end{array} \right\}. \quad (40)$$

Proof: Achievability follows immediately from Theorem 8 by extending the achievability scheme from the discrete memoryless case to the Gaussian case. The converse follows from Theorem 8 together with careful applications of the entropy power inequality and the maximum entropy theorem for Gaussian random variables. A detailed proof is given in Section V-H. ∎
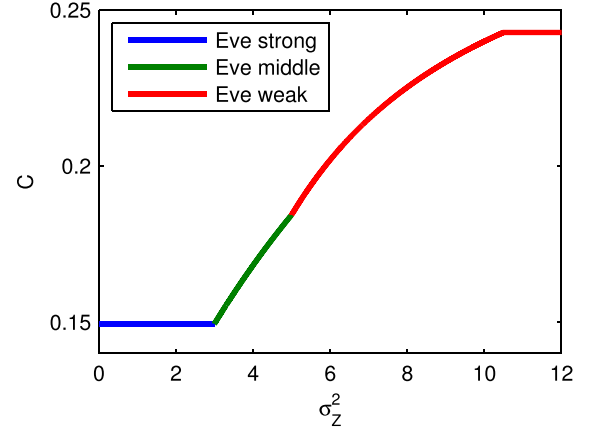

Fig. 5. Secrecy capacity of the Gaussian BC with independent secret keys for $P = 2$, $\sigma_1^2 = 3$, $\sigma_2^2 = 5$, and varying σ_Z^2.

4) Numerical Evaluation: Here we present a numerical evaluation of the secrecy capacity of the Gaussian BC with independent secret keys. Fig. 5 illustrates the behavior of the secrecy capacity for fixed noise variances of the legitimate receivers as the noise variance of the eavesdropper increases. As long as $\sigma_Z^2 < \sigma_1^2$, the eavesdropper has the strongest channel and the optimal coding scheme relies solely on the one-time pad approach. Accordingly, the secrecy capacity is constant in this regime since it depends only on the fixed legitimate channels and is independent of the actual eavesdropper channel. If $\sigma_Z^2 \geq \sigma_1^2$ but $\sigma_Z^2 < \sigma_2^2$, the eavesdropper is in the middle and the secrecy capacity is achieved by a combination of one-time pad and wiretap coding. Finally, if $\sigma_Z^2 \geq \sigma_2^2$, the eavesdropper has the weakest channel and the optimal coding scheme is wiretap coding using the secret keys as fictitious messages. At some point, the sum-rate-like bound is no longer active and the secrecy capacity is given by the capacity of the weaker link to the second receiver.

V. PROOFS

In this section we present the proofs of the main results.

A. Proof of Theorem 2

As both legitimate receivers and the eavesdropper all receive channel outputs of the same quality, we use the secret keys K_1 and K_2 as one-time pads to create secure individual messages M_1 and M_2 as in (24). Since M, K_1, and K_2 are independent of each other, we immediately have ensured that $I(M; Z^n) = 0$. Thus, the communication problem reduces to the reliable transmission of two independent individual messages. Obviously, the rate in (25) is easily achievable via time-sharing. (Note that achievability also follows immediately from the Marton coding approach in Lemma 1. By setting $U = V = W = X$ in (28) we obtain the desired rate (25). However, the time-sharing interpretation is more intuitive and yields valuable insights for the converse.)

Thus, it remains to show that time-sharing is already optimal. At receiver i, $i = 1, 2$, we have the following version of

Fano's inequality:

$$H(M|K_i, Y^n) \leq n\epsilon_{i,n} \qquad (41)$$

with $\epsilon_{i,n} \to 0$ as $n \to \infty$. Making extensive use of the definition of mutual information and the chain rule, we get

$$nR = H(M|K_i) \qquad (42)$$
$$\leq I(M; Y^n|K_i) + n\epsilon_{i,n} \qquad (43)$$
$$\leq I(M; K_i, Y^n) + n\epsilon_{i,n} \qquad (44)$$
$$= I(M; K_i|Y^n) + I(M; Y^n) + n\epsilon_{i,n} \qquad (45)$$
$$\leq I(M; K_i|Y^n) + n\epsilon_{i,n} + n\delta_n \qquad (46)$$
$$= H(K_i|Y^n) - H(K_i|M, Y^n) + n\epsilon_{i,n} + n\delta_n \qquad (47)$$
$$\leq H(K_i) - H(K_i|M, Y^n) + n\epsilon_{i,n} + n\delta_n \qquad (48)$$

where (42) follows from the independence of M and K_i, (43) from Fano's inequality (41), and (46) from the secrecy condition.

As (48) must hold for both receivers simultaneously, we obtain

$$nR \leq \min_{i \in \{1,2\}} \left\{ H(K_i) - H(K_i|M, Y^n) + n\epsilon_{i,n} + n\delta_n \right\} \qquad (49)$$

$$\leq \frac{1}{2} \Big[H(K_1) + H(K_2)$$
$$\quad - H(K_1|M, Y^n) - H(K_2|M, Y^n) + n\epsilon_n \Big] \qquad (50)$$

$$\leq \frac{1}{2} \Big[H(K_{12}) - H(K_{12}|M, Y^n) + n\epsilon_n \Big] \qquad (51)$$

$$= \frac{1}{2} \Big[H(K_{12}) + H(Y^n|M) - H(K_{12}, Y^n|M) + n\epsilon_n \Big] \qquad (52)$$

$$= \frac{1}{2} \Big[H(K_{12}) + H(Y^n|M)$$
$$\quad - H(K_{12}|M) - H(Y^n|M, K_{12}) + n\epsilon_n \Big] \qquad (53)$$

$$= \frac{1}{2} \Big[H(Y^n|M) - H(Y^n|M, K_{12}) + n\epsilon_n \Big] \qquad (54)$$

$$\leq \frac{1}{2} \Big[H(Y^n) - H(Y^n|M, K_{12}) + n\epsilon_n \Big] \qquad (55)$$

$$= \frac{1}{2} \Big[I(M, K_{12}; Y^n) + n\epsilon_n \Big] \qquad (56)$$

$$\leq \frac{1}{2} \Big[I(X^n; Y^n) + n\epsilon_n \Big] \qquad (57)$$

$$\leq \frac{1}{2} n \Big[I(X; Y) + \epsilon_n \Big] \qquad (58)$$

with $\epsilon_n = \epsilon_{1,n} + \epsilon_{2,n} + 2\delta_n$ and $\epsilon_n \to 0$ as $n \to \infty$. Here, the last step follows from Jensen's inequality and from defining the distribution P_X in (58) as the average of all P_{X_i}, $i = 1, .., n$. This completes the converse and proves the desired result. ∎

B. Proof of Theorem 3

As for the equal channel case in Theorem 2 we generate individual messages M_1 and M_2 by using the available secret keys K_1 and K_2 as one-time pads, cf. (24). Then the achievability of (27) follows immediately by superposition coding. Here, we choose the auxiliary random variable U to carry the individual message M_1 (as the "cloud center") for the weaker receiver 2. The other message M_2 for the stronger receiver 1 is superimposed as a "satellite codeword" in X. (Again, achievability can also be deduced from the Marton coding

approach in Lemma 1. By setting $U = X$ and $V = \emptyset$ in (28) we obtain the desired rate (27).)

Again, the crucial part is the converse. We have to show that this superposition coding strategy is already optimal. Using Fano's inequality (41) as in the proof of Theorem 2 led to (43) from which we proceed as follows:

$$nR \leq I(M; Y_i^n|K_i) + n\epsilon_{i,n} \qquad (59)$$
$$\leq I(M, K_i; Y_i^n) + n\epsilon_{i,n} \qquad (60)$$
$$= I(K_i; Y_i^n|M) + I(M; Y_i^n) - I(M; Z^n) + n\epsilon_n \qquad (61)$$
$$\leq I(K_i; Y_i^n|M) + n\epsilon_n \qquad (62)$$

with $\epsilon_n = \delta_n + \epsilon_{1,n}$ where (61) follows from the chain rule and the secrecy condition, and (62) from the degradedness so that $I(M; Y_i^n) - I(M; Z^n) \leq 0$, $i = 1, 2$.

Now, we define the auxiliary random variable

$$U_i := (M, K_2, Y_1^{i-1}) \qquad (63)$$

and obtain for the weaker receiver 2

$$nR \leq I(K_2; Y_2^n|M) + n\epsilon_n \qquad (64)$$
$$\leq \sum_{i=1}^n I(M, K_2; Y_{2,i}|Y_2^{i-1}) + n\epsilon_n \qquad (65)$$
$$\leq \sum_{i=1}^n I(M, K_2, Y_2^{i-1}; Y_{2,i}) + n\epsilon_n \qquad (66)$$
$$\leq \sum_{i=1}^n I(M, K_2, Y_1^{i-1}, Y_2^{i-1}; Y_{2,i}) + n\epsilon_n \qquad (67)$$
$$= \sum_{i=1}^n I(M, K_2, Y_1^{i-1}; Y_{2,i}) + n\epsilon_n \qquad (68)$$
$$= \sum_{i=1}^n I(U_i; Y_{2,i}) + n\epsilon_n \qquad (69)$$

where (68) follows from the degradedness condition. Now, let Q be a time-sharing random variable independent of all others in the model and uniformly distributed over $\{1, \ldots, n\}$. We set $U = (U_Q, Q)$, $X = X_Q$, $Y_1 = Y_{1,Q}$, and $Y_2 = Y_{2,Q}$ and obtain

$$nR \leq nI(U_Q; Y_{2,Q}|Q) + n\epsilon_n \qquad (70)$$
$$\leq nI(U; Y_2) + n\epsilon_n. \qquad (71)$$

With the same definition of U_i, cf. (63), we obtain for the stronger receiver 1

$$nR \leq I(K_1; Y_1^n|M) + n\epsilon_n \qquad (72)$$
$$\leq I(K_1; Y_1^n|M, K_2) + n\epsilon_n \qquad (73)$$
$$= \sum_{i=1}^n I(K_1; Y_{1,i}|M, K_2, Y_1^{i-1}) + n\epsilon_n \qquad (74)$$
$$\leq \sum_{i=1}^n I(K_1, X_i; Y_{1,i}|M, K_2, Y_1^{i-1}) + n\epsilon_n \qquad (75)$$
$$= \sum_{i=1}^n I(X_i; Y_{1,i}|U_i) + n\epsilon_n \qquad (76)$$
$$= nI(X_Q; Y_{1,Q}|U_Q, Q) + n\epsilon_n \qquad (77)$$
$$= nI(X; Y_1|U) + n\epsilon_n \qquad (78)$$

with $\epsilon_n = \delta_n + \epsilon_{2,n}$ which proves the converse.

The bound $|\mathcal{U}| \leq |\mathcal{X}| + 1$ on the cardinality of the range of the auxiliary random variable U follows from the strengthened version of Carathéodory's theorem, cf. for example [32], and standard arguments. The details are omitted for the sake of brevity. This completes the proof of the theorem. ∎

C. Proof of Theorem 4

The proof of the upper bound in (29) resembles ideas of [30] and [31]. Proceeding as in the proof of Theorem 3 we use Fano's inequality (41) and the fact that Y_1 and Y_2 are degraded with respect to Z to conclude on (62), i.e., at receiver i, $i = 1, 2$, the rate is upper bounded by

$$nR \leq I(K_i; Y_i^n|M) + n\epsilon_n \qquad (79)$$

with $\epsilon_n = \max\{\epsilon_{1,n}, \epsilon_{2,n}\} + \delta_n$. Now, we define the auxiliary random variables

$$U_i := (M, K_1, Y_{1,i+1}^n, Y_2^{i-1}) \qquad (80)$$
$$V_i := (M, K_2, Y_{1,i+1}^n, Y_2^{i-1}) \qquad (81)$$

and obtain for receiver 1

$$nR \leq I(K_1; Y_1^n|M) + n\epsilon_n \qquad (82)$$
$$\leq I(M, K_1; Y_1^n) + n\epsilon_n \qquad (83)$$
$$= \sum_{i=1}^{n} I(M, K_1; Y_{1,i}|Y_{1,i+1}^n) + n\epsilon_n \qquad (84)$$
$$\leq \sum_{i=1}^{n} I(M, K_1, Y_{1,i+1}^n; Y_{1,i}) + n\epsilon_n \qquad (85)$$
$$\leq \sum_{i=1}^{n} I(M, K_1, Y_{1,i+1}^n, Y_2^{i-1}; Y_{1,i}) + n\epsilon_n \qquad (86)$$
$$= \sum_{i=1}^{n} I(U_i; Y_{1,i}) + n\epsilon_n. \qquad (87)$$

Similarly, we obtain for receiver 2

$$nR \leq \sum_{i=1}^{n} I(V_i; Y_{2,i}) + n\epsilon_n. \qquad (88)$$

We continue by considering the sum of the rates in (79) of both receivers. We obtain

$$n2R \leq I(K_2; Y_2^n|M) + I(K_1; Y_1^n|M) + n\epsilon_n \qquad (89)$$
$$\leq I(K_2; K_1, Y_2^n|M) + I(M, K_1; Y_1^n) + n\epsilon_n \qquad (90)$$
$$= \sum_{i=1}^{n} I(K_2; Y_{2,i}|M, K_1, Y_2^{i-1})$$
$$+ \sum_{i=1}^{n} \left[I(M, K_1, Y_2^{i-1}; Y_{1,i}^n) - I(M, K_1, Y_2^i; Y_{1,i+1}^n) \right]$$
$$+ n\epsilon_n \qquad (91)$$

where (90) follows from the chain rule for mutual information and the last step follows the idea of the Gelfand-Pinsker converse in [33], cf. also [34, Sec. 7.9.2]

and [29, Th. 8.5 and Appendix 8B] for similar derivations. We continue as

$$n2R_n$$
$$\leq \sum_{i=1}^{n} \Big[I(K_2; Y_{2,i}|M, K_1, Y_2^{i-1}) + I(M, K_1, Y_2^{i-1}; Y_{1,i+1}^n)$$
$$+ I(M, K_1, Y_2^{i-1}; Y_{1,i}|Y_{1,i+1}^n)$$
$$- I(M, K_1, Y_2^{i-1}; Y_{1,i+1}^n)$$
$$- I(Y_{2,i}; Y_{1,i+1}^n|M, K_1, Y_2^{i-1}) \Big] + n\epsilon_n \qquad (92)$$
$$= \sum_{i=1}^{n} \Big[I(K_2; Y_{2,i}|M, K_1, Y_2^{i-1})$$
$$+ I(M, K_1, Y_2^{i-1}; Y_{1,i}|Y_{1,i+1}^n)$$
$$- I(Y_{2,i}; Y_{1,i+1}^n|M, K_1, Y_2^{i-1}) \Big] + n\epsilon_n \qquad (93)$$
$$= \sum_{i=1}^{n} \Big[H(Y_{2,i}|M, K_1, Y_2^{i-1}) - H(Y_{2,i}|M, K_1, K_2, Y_2^{i-1})$$
$$+ H(Y_{1,i}|Y_{1,i+1}^n) - H(Y_{1,i}|M, K_1, Y_{1,i+1}^n, Y_2^{i-1})$$
$$- H(Y_{2,i}|M, K_1, Y_2^{i-1}) + H(Y_{2,i}|M, K_1, Y_{1,i+1}^n, Y_2^{i-1}) \Big]$$
$$+ n\epsilon_n \qquad (94)$$
$$\leq \sum_{i=1}^{n} \Big[H(Y_{1,i}) - H(Y_{1,i}|U_i) + H(Y_{2,i}|U_i) - H(Y_{2,i}|K_2, U_i) \Big]$$
$$+ n\epsilon_n \qquad (95)$$
$$\leq \sum_{i=1}^{n} \Big[I(U_i; Y_{1,i}) + H(Y_{2,i}|U_i) - H(Y_{2,i}|U_i, X_i) \Big] + n\epsilon_n \qquad (96)$$
$$\leq \sum_{i=1}^{n} \Big[I(U_i; Y_{1,i}) + I(X_i; Y_{2,i}|U_i) \Big] + n\epsilon_n. \qquad (97)$$

Similarly, by interchanging the order of Y_1 and Y_2 we obtain

$$n2R \leq I(K_1; Y_1^n|M) + I(K_2; Y_2^n|M) + n\epsilon_n \qquad (98)$$
$$\leq I(K_1; K_2, Y_1^n|M) + I(M, K_2; Y_2^n) + n\epsilon_n \qquad (99)$$
$$= \sum_{i=1}^{n} I(K_1; Y_{1,i}|M, K_2, Y_{1,i+1}^n)$$
$$+ \sum_{i=1}^{n} \Big[I(M, K_2, Y_{1,i+1}^n; Y_2^i) - I(M, K_2, Y_{1,i}^n; Y_2^{i-1}) \Big]$$
$$+ n\epsilon_n \qquad (100)$$

where the last step follows the idea of the Gelfand-Pinsker converse, cf. [33]. We continue as

$$n2R \leq \sum_{i=1}^{n} \Big[I(K_1; Y_{1,i}|M, K_2, Y_{1,i+1}^n)$$
$$+ I(M, K_2, Y_{1,i+1}^n; Y_2^{i-1}) + I(M, K_2, Y_{1,i+1}^n; Y_{2,i}|Y_2^{i-1})$$
$$- I(M, K_2, Y_{1,i+1}^n; Y_2^{i-1}) - I(Y_{1,i}; Y_2^{i-1}|M, K_2, Y_{1,i+1}^n) \Big]$$
$$+ n\epsilon_n \qquad (101)$$
$$= \sum_{i=1}^{n} \Big[I(K_1; Y_{1,i}|M, K_2, Y_{1,i+1}^n)$$
$$+ I(M, K_2, Y_{1,i+1}^n; Y_{2,i}|Y_2^{i-1})$$
$$- I(Y_{1,i}; Y_2^{i-1}|M, K_2, Y_{1,i+1}^n) \Big] + n\epsilon_n \qquad (102)$$

$$\begin{aligned}
&= \sum_{i=1}^{n} \Big[H(Y_{1,i}|M, K_2, Y_{1,i+1}^n) - H(Y_{1,i}|M, K_1, K_2, Y_{1,i+1}^n) \\
&\quad + H(Y_{2,i}|Y_2^{i-1}) - H(Y_{2,i}|M, K_2, Y_{1,i+1}^n, Y_2^{i-1}) \\
&\quad - H(Y_{1,i}|M, K_2, Y_{1,i+1}^n) \\
&\quad + H(Y_{1,i}|M, K_2, Y_{1,i+1}^n, Y_2^{i-1}) \Big] + n\epsilon_n
\end{aligned} \tag{103}$$

$$\begin{aligned}
&\leq \sum_{i=1}^{n} \Big[H(Y_{2,i}) - H(Y_{2,i}|V_i) \\
&\quad + H(Y_{1,i}|V_i) - H(Y_{1,i}|K_1, V_i) \Big] + n\epsilon_n
\end{aligned} \tag{104}$$

$$\leq \sum_{i=1}^{n} \Big[I(V_i; Y_{2,i}) + H(Y_{1,i}|V_i) - H(Y_{1,i}|V_i, X_i) \Big] + n\epsilon_n \tag{105}$$

$$\leq \sum_{i=1}^{n} \Big[I(V_i; Y_{2,i}) + I(X_i; Y_{1,i}|V_i) \Big] + n\epsilon_n. \tag{106}$$

Now, let Q be a time-sharing random variable independent of all others in the model and uniformly distributed over $\{1, \dots, n\}$. We set $U = (U_Q, Q)$, $V = (V_Q, Q)$, $X = X_Q$, $Y_1 = Y_{1,Q}$, and $Y_2 = Y_{2,Q}$ and obtain for (87)

$$nR \leq \sum_{i=1}^{n} I(U_i; Y_{1,i}) + n\epsilon_n \tag{107}$$

$$= nI(U_Q; Y_{1,Q}|Q) + n\epsilon_n \tag{108}$$

$$\leq nI(U; Y_1) + n\epsilon_n \tag{109}$$

and similarly for (88)

$$nR \leq nI(V; Y_2) + n\epsilon_n. \tag{110}$$

Accordingly, we get for (97) and (106)

$$n2R \leq \sum_{i=1}^{n} \Big[I(U_i; Y_{1,i}) + I(X_i; Y_{2,i}|U_i) \Big] + n\epsilon_n \tag{111}$$

$$= n\Big[I(U_Q; Y_{1,Q}|Q) + I(X_Q; Y_{2,Q}|U_Q, Q) \Big] + n\epsilon_n \tag{112}$$

$$\leq n\Big[I(U; Y_1) + I(X; Y_2|U) \Big] + n\epsilon_n \tag{113}$$

and

$$n2R \leq n\Big[I(V; Y_2) + I(X; Y_1|V) \Big] + n\epsilon_n. \tag{114}$$

This establishes the desired bounds in (29) and it remains to bound the cardinalities as stated. This can be done exactly as for the UV-bound for the classical BC in [35, Sec. II-C] and is therefore omitted for brevity. This completes the proof. ∎

D. Proof of Theorem 5

Achievability follows from the Marton coding approach in Lemma 1. By setting $U = Y_1$, $V = Y_2$, and $W = \emptyset$, the rate expression (28) simplifies to the desired rate (30).

The converse follows then from the UV-type upper bound in Theorem 4. We observe that the first two bounds in (29)

can be bounded as $I(U; Y_1) \leq H(Y_1)$ and $I(V; Y_2) \leq H(Y_2)$. Further,

$$\begin{aligned}
&I(U; Y_1) + I(X; Y_2|U) \\
&\quad = I(U; Y_1) + H(Y_2|U) \tag{115} \\
&\quad = H(Y_1) - H(Y_1|U) + H(Y_2|U) \tag{116} \\
&\quad = H(Y_1) + H(Y_2|Y_1) \\
&\qquad - H(Y_2|Y_1) - H(Y_1|U) + H(Y_2|U) \tag{117} \\
&\quad \leq H(Y_1, Y_2) - H(Y_2|Y_1, U) - H(Y_1|U) + H(Y_2|U) \tag{118} \\
&\quad = H(Y_1, Y_2) - H(Y_1, Y_2|U) + H(Y_2|U) \tag{119} \\
&\quad \leq H(Y_1, Y_2) \tag{120}
\end{aligned}$$

where (115) follows from $H(Y_2|X, U) = 0$ since the channel is deterministic and (120) follows from the fact that $-H(Y_1, Y_2|U) + H(Y_2|U) \leq 0$. In a similar way we also get $\frac{1}{2}[I(V; Y_2) + I(X; Y_1|V)] \leq H(Y_1, Y_2)$ which give the desired bounds. ∎

E. Proof of Theorem 6

We start with the achievability of the rate (31) and then prove the corresponding converse.

1) Proof of Achievability: The following equivalent description of (31) turns out to be beneficial for the proof of achievability.

Lemma 2: The rate expression in (31) can equivalently be expressed as

$$C = \max_{P_X} \max_{0 \leq \alpha \leq 1} \min \left\{ \begin{matrix} I(X; Y_1) - \alpha I(X; Z) \\ I(X; Y_2) - (1 - \alpha)I(X; Z) \end{matrix} \right\}. \tag{121}$$

Proof: A proof is provided in Appendix B. ∎

Thus, instead of proving the achievability of (31), we prove the achievability of (121) for any $0 \leq \alpha \leq 1$.

Next we sketch the proof of achievability. Basically, it follows the ideas of [27], [28], and [36] which all present coding schemes that achieve strong secrecy as required in (23). Accordingly, for any input distribution $P_X \in \mathcal{P}(\mathcal{X})$ and $\alpha \in [0, 1]$ we generate $|\mathcal{M}||\mathcal{K}_1||\mathcal{K}_2|$ independent codewords $x_{mk_1k_2}^n \in \mathcal{X}^n$ where

$$|\mathcal{K}_1| > 2^{n((1-\alpha)I(X;Z)+\epsilon)} \tag{122a}$$

$$|\mathcal{K}_2| > 2^{n(\alpha I(X;Z)+\epsilon)} \tag{122b}$$

$$|\mathcal{M}| < \min \left\{ \begin{matrix} 2^{n(I(X;Y_1)-\alpha I(X;Z)-2\epsilon)} \\ 2^{n(I(X;Y_2)-(1-\alpha)I(X;Z)-2\epsilon)} \end{matrix} \right\}. \tag{122c}$$

The crucial idea is to use the available secret keys as randomization resources instead of generating "dummy" randomization indices as in the classical wiretap coding approach. As the sizes of the secret keys satisfy

$$\frac{1}{n} \log(|\mathcal{K}_1||\mathcal{K}_2|) > I(X; Z) + 2\epsilon \tag{123}$$

we have enough randomization resources to show that $I(M; Z^n) \leq \delta_n$ holds, i.e., strong secrecy (23) is satisfied. This can be done similarly as in [27], [28], and [36].

Next, we check the reliability constraints at the legitimate receivers. Receiver 1 has the secret key $k_1 \in \mathcal{K}_1$ as side information available and therefore the unknown indices of

the transmitted codeword are $m \in \mathcal{M}$ and $k_2 \in \mathcal{K}_2$. As its size satisfies

$$|\mathcal{M}||\mathcal{K}_2| \leq 2^{n(I(X;Y_1)-\epsilon)}, \qquad (124)$$

it is straightforward to show that receiver 1 can decode the remaining indices $m \in \mathcal{M}$ and $k_2 \in \mathcal{K}_2$. Similarly, receiver 2 has $k_2 \in \mathcal{K}_2$ as side information available and the unknown indices are $m \in \mathcal{M}$ and $k_1 \in \mathcal{K}_1$ of size

$$|\mathcal{M}||\mathcal{K}_1| \leq 2^{n(I(X;Y_2)-\epsilon)}. \qquad (125)$$

Again, it is easy to show that receiver 2 can decode the remaining indices $m \in \mathcal{M}$ and $k_1 \in \mathcal{K}_1$. Thus, we conclude that (121) is an achievable rate.

Remark 12: For classical wiretap coding, the rate needed for additional randomization is roughly $I(X; Z)$. This suffices to keep the eavesdropper ignorant of the transmitted message. The use of secret keys as the randomization resource has the advantage that parts of the needed randomization resource are already available as side information at the receivers. This reduces the loss in rate in the sense that it is only reduced by the remaining unknown randomization part (and not by the whole amount of randomization).

2) Proof of the Converse: It remains to show the optimality of the above presented coding scheme. The first two bounds in (31) are the obvious single-user bounds and follow immediately. The crucial part is to prove the third "sum-rate"-like bound. We proceed as follows:

$$n2R \leq H(M) + H(M) \qquad (126)$$
$$= H(M|K_1) + H(M|K_2) \qquad (127)$$
$$\leq I(M; Y_1^n|K_1) + I(M; Y_2^n|K_2) + n\epsilon_{1,n} + n\epsilon_{2,n} \qquad (128)$$
$$\leq I(M; Y_1^n|K_1) + I(M; Y_2^n|K_2)$$
$$\quad - I(M; Z^n) + n\epsilon_n \qquad (129)$$
$$\leq I(M, K_1; Y_1^n) + I(M, K_2; Y_2^n)$$
$$\quad - I(M; Z^n) + n\epsilon_n \qquad (130)$$
$$= I(M, K_{12}; Y_1^n) + I(M, K_{12}; Y_2^n)$$
$$\quad - I(M, K_{12}; Z^n) - I(K_2; Y_1^n|M, K_1)$$
$$\quad - I(K_1; Y_2^n|M, K_2) + I(K_{12}; Z^n|M) + n\epsilon_n \qquad (131)$$
$$\leq I(M, K_{12}; Y_1^n) + I(M, K_{12}; Y_2^n)$$
$$\quad - I(M, K_{12}; Z^n) + n\epsilon_n \qquad (132)$$

with $\epsilon_n = \delta_n + \epsilon_{1,n} + \epsilon_{2,n}$ and $\epsilon_n \to 0$ as $n \to \infty$. Here, (128) follows from Fano's inequality, cf. (41), (129) from the secrecy criterion, and (132) from the fact that $-I(K_2; Y_1^n|M, K_1) - I(K_1; Y_2^n|M, K_2) + I(K_1, K_2; Z^n|M) \leq 0$. To see this last step, we write

$$-I(K_2; Y_1^n|M, K_1) - I(K_1; Y_2^n|M, K_2) + I(K_{12}; Z^n|M)$$
$$= -H(K_2|M, K_1) + H(K_2|M, K_1, Y_1^n)$$
$$\quad - H(K_1|M, K_2) + H(K_1|M, K_2, Y_2^n)$$
$$\quad + H(K_{12}|M) - H(K_{12}|M, Z^n) \qquad (133)$$
$$= H(K_2|M, K_1, Y_1^n) + H(K_1|M, K_2, Y_2^n)$$
$$\quad - H(K_{12}|M, Z^n) \qquad (134)$$
$$\leq H(K_2|M, K_1, Z^n) + H(K_1|M, K_2, Z^n)$$
$$\quad - H(K_{12}|M, Z^n) \qquad (135)$$
$$\leq 0 \qquad (136)$$

where (134) follows from the fact that M, K_1, and K_2 are independent so that $-H(K_2|M, K_1) - H(K_1|M, K_2) + H(K_{12}|M) = 0$, and (135) from the Markov chain relationships $X - Y_1 - Z$ and $X - Y_2 - Z$ due to the degradedness. Now, with this we can proceed with the "sum-rate" in (132) as

$$n2R \leq I(M, K_{12}; Y_1^n) + I(M, K_{12}; Y_2^n)$$
$$\quad - I(M, K_{12}; Z^n) + n\epsilon_n \qquad (137)$$
$$= I(M, K_{12}; Y_1^n|Z^n) + I(M, K_{12}; Y_2^n) + n\epsilon_n \qquad (138)$$
$$\leq I(X^n; Y_1^n|Z^n) + I(X^n; Y_2^n) + n\epsilon_n \qquad (139)$$
$$\leq n[I(X; Y_1|Z) + I(X; Y_2)] + n\epsilon_n \qquad (140)$$
$$= n[I(X; Y_1) + I(X; Y_2) - I(X; Z)] + n\epsilon_n \qquad (141)$$

where (138) and (141) follow from the degradedness of the channels. This completes the proof of the converse. ∎

F. Proof of Theorem 7

The proof of achievability follows immediately from Theorem 6 so that it remains to prove the converse for the less noisy case. In the following we will highlight the differences from the converse proof for the degraded case in Theorem 6.

Following the lines (126)-(132) for the "sum-rate" term we also obtain

$$n2R \leq I(M, K_{12}; Y_1^n) + I(M, K_{12}; Y_2^n)$$
$$\quad - I(M, K_{12}; Z^n) - I(K_2; Y_1^n|M, K_1)$$
$$\quad - I(K_1; Y_2^n|M, K_2) + I(K_{12}; Z^n|M) + n\epsilon_n \qquad (142)$$
$$\leq I(M, K_{12}; Y_1^n) + I(M, K_{12}; Y_2^n)$$
$$\quad - I(M, K_{12}; Z^n) + n\epsilon_n \qquad (143)$$

with $\epsilon_n = \delta_n + \epsilon_{1,n} + \epsilon_{2,n}$ and $\epsilon_n \to 0$ as $n \to \infty$. Here, (143) follows again from the fact that $-I(K_2; Y_1^n|M, K_1) - I(K_1; Y_2^n|M, K_2) + I(K_1, K_2; Z^n|M) \leq 0$. To see that this also holds for less noisy channels, we proceed as follows. With $I(K_1, K_2; Z^n|M) = I(K_1; Z^n|M) + I(K_2; Z^n|M, K_1)$ we observe that

$$-I(K_2; Y_1^n|M, K_1) + I(K_2; Z^n|M, K_1) \leq 0 \qquad (144)$$

which is a consequence of the less noisy condition $Y_1 \succeq Z$, cf. for example [37, Proposition 1]. With $I(K_1; Z^n|M) \leq I(K_1; Z^n|M, K_2)$ we also have

$$-I(K_1; Y_2^n|M, K_2) + I(K_1; Z^n|M, K_2) \leq 0 \qquad (145)$$

which validates the correctness of the inequality in (143).

Now, we can proceed with the "sum-rate" in (143) as

$$n2R$$
$$\leq I(M, K_{12}; Y_1^n) + I(M, K_{12}; Y_2^n)$$
$$\quad - I(M, K_{12}; Z^n) + n\epsilon_n \qquad (146)$$
$$\leq \sum_{i=1}^{n} \left[I(M, K_{12}; Y_{1,i}|Y_1^{i-1}) - I(M, K_{12}; Z_i|Z_{i+1}^n) \right]$$
$$\quad + I(X^n; Y_2^n) + n\epsilon_n \qquad (147)$$

$$= \sum_{i=1}^{n} \Big[I(M, K_{12}; Y_{1,i}|Y_1^{i-1}, Z_{i+1}^n)$$

$$- I(M, K_{12}; Z_i|Y_1^{i-1}, Z_{i+1}^n) \Big] + I(X^n; Y_2^n) + n\epsilon_n \quad (148)$$

$$= \sum_{i=1}^{n} \Big[I(V_i; Y_{1,i}|U_i) - I(V_i : Z_i|U_i) \Big] + I(X^n; Y_2^n) + n\epsilon_n$$
$$(149)$$

where (148) follows from the Csiszár sum identity [10, Lemma 7] and (149) from the definition of auxiliary random variables $U_i := (Y_1^{i-1}, Z_{i+1}^n)$ and $V_i := (M, K_{12}, U_i)$. Now, let Q be a time-sharing random variable independent of all others in the model and uniformly distributed over $\{1, \ldots, n\}$. We set $U = (U_Q, Q)$, $V = V_Q$, $X = X_Q$, $Y_1 = Y_{1,Q}$, and $Z = Z_Q$ to obtain

$$n2R \leq n \Big[I(V; Y_1|U) - I(V; Z|U) + I(X; Y_2) \Big] + n\epsilon_n$$
$$(150)$$

$$\leq n \Big[I(V^*; Y_1) - I(V^*; Z) + I(X; Y_2) \Big] + n\epsilon_n \quad (151)$$

$$= n \Big[I(X; Y_1) - I(X; Y_1|V^*)$$
$$- I(X; Z) + I(X; Z|V^*) + I(X; Y_2) \Big] + n\epsilon_n \quad (152)$$

$$\leq n \Big[I(X; Y_1) + I(X; Y_2) - I(X; Z) \Big] + n\epsilon_n \quad (153)$$

where (151) follows from letting V^* be a random variable distributed according to the choice $U = u^*$ that maximizes the difference in (150). Finally, the inequality in (153) follows from the less noisy condition $Y \succeq Z$ which implies $I(X; Y_1|V^*) \geq I(X; Z|V^*)$. This concludes the proof of the converse. ∎

G. Proof of Theorem 8

Proof of Achievability: For achievability we want to exploit both previously discussed approaches: secret keys as one-time pads and secret keys as fictitious messages in wiretap codes. Therefore, we split the message M into two parts

$$M = (M', M^{\oplus}) \quad (154)$$

with rates R' and $R^{\oplus}$, where the part M' will be protected by the wiretap coding approach and $M^{\oplus}$ by the one-time pad approach. Accordingly, we split the secret keys as

$$K_1 = (K_1', K_1^{\oplus}) \quad \text{and} K_2 = (K_2', K_2^{\oplus}) \quad (155)$$

such that $R^{\oplus} = R_{K_1}^{\oplus} = R_{K_2}^{\oplus}$ to ensure that $M^{\oplus}$ can be protected following the one-time pad idea by creating encrypted messages

$$M_1^{\oplus} = M^{\oplus} \oplus K_1^{\oplus} \quad \text{and} \quad M_2^{\oplus} = M^{\oplus} \oplus K_2^{\oplus} \quad (156)$$

both of rate $R^{\oplus}$.

We use a superposition coding scheme with two layers. For the random code construction, for any P_U we first generate "cloud centers" carrying M', K_1', K_2', and $M_2^{\oplus}$. Here, M' will be protected by a wiretap code approach. Then, according to $P_{X|U}$ we generate "satellite codewords" carrying $M_1^{\oplus}$.

Now, the stronger receiver 1 decodes the cloud center (having K_1' as side information available) and the satellite codeword, while the weaker receiver 2 decodes only the cloud

center (having K_2' as side information available). If we choose the rates of all messages such that they satisfy

$$R' + R_{K_2}' + R^{\oplus} \leq I(U; Y_1) \quad (157)$$
$$R' + R_{K_1}' + R^{\oplus} \leq I(U; Y_2) \quad (158)$$
$$R^{\oplus} \leq I(X; Y_1|U) \quad (159)$$
$$R_{K_1}' + R_{K_2}' + R^{\oplus} > I(U; Z) \quad (160)$$
$$R^{\oplus} > I(X; Z|U) \quad (161)$$

the average probabilities of error become arbitrarily small. Having decoded $M_1^{\oplus}$ and $M_2^{\oplus}$ respectively, the receivers obtain the desired part $M^{\oplus}$ with the help of their side information $K_1^{\oplus}$ and $K_2^{\oplus}$ respectively.

It remains to check the security of the transmitted message. As $R_{K_1}' + R_{K_2}' + R^{\oplus} > I(U; Z)$ and $R^{\oplus} > I(X; Z|U)$, the rate of the fictitious messages is high enough to show that $I(M'; Z^n) \leq \delta_n$ holds, i.e., strong secrecy (23) is satisfied. This can be done similarly as in [27], [28], and [36]. Note that no information about $M^{\oplus}$ is leaked to the eavesdropper due to the one-time pad $K_2^{\oplus}$ even if the eavesdropper is able to decode the encrypted message $M_2^{\oplus}$. The second part $M_1^{\oplus}$ in the satellite codeword is immediately protected by the one-time pad $K_1^{\oplus}$.

Now, by applying Fourier-Motzkin elimination to get rid of R', $R^{\oplus}$, R_{K_1}', and R_{K_2}' in (157)-(161) having in mind that $R = R' + R^{\oplus}$, yields

$$R \leq \min \left\{ \begin{array}{l} I(U; Y_1), \\ I(U; Y_2), \\ \frac{1}{2} \big[I(X; Y_1) + I(U; Y_2) - I(U; Z) \big] \end{array} \right\} \quad (162)$$

where we observe that the first condition can be omitted as $I(U; Y_2) \leq I(U; Y_1)$ holds due to the degradedness. This concludes the proof of achievability.

Remark 13: We want to highlight the difference from the previous approach in Theorem 6. There, the available secret keys are solely used as fictitious messages. From a legitimate receiver's perspective, using its own secret key as a fictitious message increases the equivocation rate at the eavesdropper but does not contribute to the message rate. However, Theorem 6 shows that for degraded channels $X - Y_1 - Z$ and $X - Y_2 - Z$ this is already sufficient to achieve capacity, cf. Theorem 6.

The one-time pad approach enters the picture when one legitimate receiver is stronger than the eavesdropper. Additionally to using the secret keys as fictitious messages, one part of the secret key is used to create an encrypted message $M_2^{\oplus}$ which is also used as a fictitious message part to increase the equivocation rate. Again, it does not contribute to the message rate of receiver 1, but it is useful for receiver 2. To resolve the mismatch, the stronger receiver 1 receives the missing part encrypted in the satellite codeword.

This approach is in particular beneficial if receiver 2 is much weaker than the eavesdropper while receiver 1 and the eavesdropper are close in terms of channel quality.

Proof of the Converse: We define the auxiliary random variable

$$U_i := (M, K_2, Z^{i-1}). \quad (163)$$

The first bound in (34) is a single-user bound and follows as

$$nR \leq I(M; Y_2^n | K_2) + n\epsilon_{1,n} \tag{164}$$

$$\leq I(M, K_2; Y_2^n) + n\epsilon_{1,n} \tag{165}$$

$$\leq \sum_{i=1}^{n} I(M, K_2; Y_{2,i} | Y_2^{i-1}) + n\epsilon_{1,n} \tag{166}$$

$$\leq \sum_{i=1}^{n} I(M, K_2, Y_2^{i-1}; Y_{2,i}) + n\epsilon_{1,n} \tag{167}$$

$$\leq \sum_{i=1}^{n} I(M, K_2, Y_2^{i-1}, Z^{i-1}; Y_{2,i}) + n\epsilon_{1,n} \tag{168}$$

$$\leq \sum_{i=1}^{n} I(M, K_2, Z^{i-1}; Y_{2,i}) + n\epsilon_{1,n} \tag{169}$$

$$= \sum_{i=1}^{n} I(U_i; Y_{2,i}) + n\epsilon_{1,n} \tag{170}$$

with $\epsilon_{1,n} \to 0$ as $n \to \infty$. Here, (164) follows from Fano's inequality and (170) from the definition of U_i, cf. (163).

The crucial part is the second "sum-rate"-like bound. From the proof of [1, Th. 6], we know that $n2R \leq I(M, K_1; Y_1^n) + I(M, K_2; Y_2^n) - I(M; Z^n) + n\epsilon_n$. Following this, we obtain

$$n2R \leq I(M, K_1; Y_1^n) + I(M, K_2; Y_2^n)$$
$$\quad - I(M; Z^n) + n\epsilon_n \tag{171}$$

$$= I(M, K_{12}; Y_1^n) + I(M, K_2; Y_2^n) - I(M, K_2; Z^n)$$
$$\quad - I(K_2; Y_1^n | M, K_1) + I(K_2; Z^n | M) + n\epsilon_n \tag{172}$$

$$\leq I(M, K_{12}; Y_1^n) + I(M, K_2; Y_2^n)$$
$$\quad - I(M, K_2; Z^n) + n\epsilon_n \tag{173}$$

with $\epsilon_n = \delta_n + \epsilon_{1,n} + \epsilon_{2,n}$ and $\epsilon_n \to 0$ as $n \to \infty$. Here, (173) follows from the fact that $-I(K_2; Y_1^n | M, K_1) + I(K_2; Z^n | M) \leq 0$. Now we proceed as

$$n2R \leq I(M, K_{12}; Y_1^n) + I(M, K_2; Y_2^n)$$
$$\quad - I(M, K_2; Z^n) + n\epsilon_n \tag{174}$$

$$= I(M, K_{12}; Y_1^n) + I(M, K_2; Y_2^n) - I(M, K_{12}; Z^n)$$
$$\quad + I(K_1; Z^n | M, K_2) + n\epsilon_n \tag{175}$$

where the first and third term are bounded as

$$I(M, K_{12}; Y_1^n) - I(M, K_{12}; Z^n)$$
$$= I(M, K_{12}; Y_1^n | Z^n) \leq I(X^n; Y_1^n | Z^n) \tag{176}$$

$$\leq n[I(X; Y_1 | Z)] = n[I(X; Y_1) - I(X; Z)]. \tag{177}$$

Here, (176) holds due $X - Y_1 - Z$. The second term becomes

$$I(M, K_2; Y_2^n) \leq \sum_{i=1}^{n} I(U_i; Y_{2,i}) \tag{178}$$

which follows as in (165)-(170). And finally, the last term is

$$I(K_1; Z^n | M, K_2) = \sum_{i=1}^{n} I(K_1; Z_i | M, K_2, Z^{i-1}) \tag{179}$$

$$\leq \sum_{i=1}^{n} I(K_1, X_i; Z_i | M, K_2, Z^{i-1}) \tag{180}$$

$$= \sum_{i=1}^{n} I(X_i; Z_i | M, K_2, Z^{i-1})$$

$$= \sum_{i=1}^{n} I(X_i; Z_i | U_i). \tag{181}$$

Now, by following standard arguments of introducing a time-sharing random variable independent of all others and uniformly distributed over $\{1, \ldots, n\}$, it is straightforward to finally obtain $R \leq I(U; Y_2) + \epsilon_{1,n}$ and $2R \leq [I(X; Y_1) + I(U; Y_2) - I(U; Z)] + \epsilon_n$.

The bound $|\mathcal{U}| \leq |\mathcal{X}| + 1$ on the cardinality of the range of the auxiliary random variable U follows from the strengthened version of Carathéodory's theorem, cf. [32]. ∎

H. Proof of Theorem 11

Achievability follows from the corresponding scheme for discrete memoryless channels in Theorem 8, where we choose $U \sim \alpha P$ (wiretap coding), $V \sim (1-\alpha)P$ (additional one-time pad part), and $X = U + V$ with $0 \leq \alpha \leq 1$ parametrizing the power allocation between the two schemes.

The converse follows from an application of the entropy power inequality and the maximum entropy theorem. We write the first condition in (34) as

$$I(U; Y_2) = h(Y_2) - h(Y_2 | U) \tag{182}$$

with $h(\cdot)$ the differential entropy. By the maximum entropy theorem, the first term is upper bounded by $h(Y_2) \leq \frac{1}{2} \log(2\pi e(P + \sigma_2^2))$. For the second term $h(Y_2 | U)$ we observe

$$\frac{1}{2} \log(2\pi e \sigma_2^2) = h(N_2) = h(Y_2 | X) \leq h(Y_2 | U) \tag{183}$$

$$\leq h(Y_2) \leq \frac{1}{2} \log(2\pi e(P + \sigma_2^2)) \tag{184}$$

so that there must exist an $\alpha \in [0, 1]$ such that

$$h(Y_2 | U) = \frac{1}{2} \log \left(2\pi e((1 - \alpha)P + \sigma_2^2)\right). \tag{185}$$

With this, we obtain for the first expression in (34)

$$I(U; Y_2) \leq \frac{1}{2} \log \left(1 + \frac{\alpha P}{(1 - \alpha)P + \sigma_2^2}\right) \tag{186}$$

which already establishes the first bound in (40) as desired.

With $I(X; Z) = I(U; Z) + I(X; Z | U)$ we can write the second condition in (34) as

$$I(X; Y_1) + I(U; Y_2) - I(X; Z) + I(X; Z | U). \tag{187}$$

In the following, we will bound each term separately. From (186) we already know that $I(U; Y_2) \leq \frac{1}{2} \log(1 + \frac{\alpha P}{(1-\alpha)P + \sigma_2^2})$. Following the proof for the classical Gaussian wiretap channel by using the entropy power inequality, we further get

$$I(X; Y_1) - I(X; Z) \leq \frac{1}{2} \log \left(1 + \frac{P}{\sigma_1^2}\right) - \frac{1}{2} \log \left(1 + \frac{P}{\sigma_Z^2}\right). \tag{188}$$

Thus, it remains to bound the last term $I(X; Z | U)$ as

$$I(X; Z | U) = h(Z | U) - h(Z | U, X) \tag{189}$$

$$= h(Z | U) - \frac{1}{2} \log(2\pi e \sigma_Z^2). \tag{190}$$

To bound the remaining term, we apply the entropy power inequality as follows:

$$h(Y_2|U) = h(Z + \tilde{N}|U) \tag{191}$$

$$\geq \frac{1}{2} \log \left(2^{2h(Z|U)} + 2^{2h(\tilde{N}|U)} \right) \tag{192}$$

$$= \frac{1}{2} \log \left(2^{2h(Z|U)} + 2\pi e(\sigma_2^2 - \sigma_Z^2) \right) \tag{193}$$

where $\tilde{N}$ is Gaussian distributed with variance $\sigma_2^2 - \sigma_Z^2$. Now, we choose α as in (185) to obtain

$$2\pi e((1-\alpha)P + \sigma_2^2) \geq 2^{2h(Z|U)} + 2\pi e(\sigma_2^2 - \sigma_Z^2) \tag{194}$$

so that

$$h(Z|U) \leq \frac{1}{2} \log \left(2\pi e((1-\alpha)P + \sigma_Z^2) \right). \tag{195}$$

Now, putting everything together yields

$$I(X; Y_1) + I(U; Y_2) - I(U; Z)$$
$$\leq \frac{1}{2} \Big[\log \left(1 + \frac{P}{\sigma_1^2} \right) + \log \left(1 + \frac{\alpha P}{(1-\alpha)P + \sigma_2^2} \right)$$
$$- \log \left(1 + \frac{\alpha P}{(1-\alpha)P + \sigma_Z^2} \right) \Big] \tag{196}$$

which yields the remaining bound in (40). This completes the converse and therewith the proof. ∎

VI. CONCLUSION

A secret key shared between a transmitter and a receiver clearly helps to securely communicate in the presence of an eavesdropper. In the "single-user" user scenario of the wiretap channel with shared secret key, the gain in secrecy capacity can immediately be quantified. It corresponds to the key rate of the shared secret key which agrees with common intuition. The same question becomes much more involved and non-trivial to answer for more complex communication scenarios in which multiple secret keys are shared among the legitimate users. Shared secret keys suggest the use of one-time pads to encrypt confidential messages for keeping external eavesdroppers ignorant. However, a secret key shared between the transmitter and one receiver might harm other receivers which do not share this key. Thus, multiple secret keys can result in conflicting payoffs at different receivers, which raises the question of how these keys should be used in an optimal way. In this paper, we have studied the BC with independent secret keys in which a common message must be securely transmitted to two legitimate receivers each of which shares an independent secret key with the transmitter. This is the simplest extension beyond the single-user wiretap channel that captures the above mentioned effects.

For situations in which the eavesdropper's channel is "stronger" than the legitimate channels, classical wiretap coding does not work and the message must be protected by using the secret keys as one-time pads. For several special cases such as the noiseless BC, BC with equal channel outputs, or degraded BC, it has been shown that this strategy is actually optimal and the secrecy capacity has been established.

On the other hand, in situations in which the eavesdropper's channel is the "weakest" channel, it has been shown that it is optimal to use the secret keys not as one-time pads but as fictitious messages for randomization in wiretap coding. The corresponding secrecy capacity has been established as well. The most challenging case is when the eavesdropper has neither the strongest nor the weakest channel. Here, a careful combination of the one-time pad approach and wiretap coding with fictitious messages is needed to establish the secrecy capacity for this case. First, these problems have been addressed for the situation in which all channels are discrete memoryless. Then, Gaussian channels have been considered and the previous results have been extended to this case as well yielding a complete characterization of the secrecy capacity of the Gaussian BC with independent secret keys.

For future work, it would be interesting to study how these results obtained for the case of two legitimate receivers extend to an arbitrary number of legitimate receivers. Another direction for future work would be to characterize how large of secret key rates are needed to achieve capacity and what happens if the secret key rates are bounded below those rates.

APPENDIX

A. Sketch of Proof of Lemma 1

As described in Section IV-A, the secret keys are used as one-time pads to create individual encrypted messages $M_1 = M \oplus K_1$ and $M_2 = M \oplus K_2$, cf. (24) and Fig. 3, which are then encoded and transmitted. As M, K_1, and K_2 are independent of each other, we immediately have $I(M; Z^n) = 0$ even if the eavesdropper would be able to decode the encrypted messages M_1 or M_2.

Now, the achievability of the rate given in (28) follows from the idea of Marton coding for the classical BC, cf. [29, Sec. 8.4]. In particular, the created individual messages M_1 and M_2 are encoded following this approach which yields the bounds on the rate as given in (28), cf. [29, Proposition 8.1]. ∎

B. Proof of Lemma 2

With the function

$$f(t) = \begin{cases} (1-t)I(X; Y_1) + t[I(X; Y_2) - I(X; Z)] & \text{if } t \leq \frac{1}{2} \\ (1-t)[I(X; Y_1) - I(X; Z)] + tI(X; Y_2) & \text{if } t \geq \frac{1}{2} \end{cases} \tag{197}$$

we can express the rate expression in (31) as

$$C = \max_{P_X} \min_{0 \leq t \leq 1} f(t). \tag{198}$$

As the function $f(t)$ is piecewise linear, it is sufficient to evaluate it at the corner points, i.e., when $t = 0$, $t = 1$, and $t = \frac{1}{2}$, to convince ourself that the rate expressions in (31) and (198) are equivalent.

Now we have to show that (198) is equivalent to the desired expression (121). Therefore, we rewrite (121) (where we omit

the outer maximization for short) as

$$\max_{0 \le \alpha \le 1} \min_{0 \le t \le 1} \Big[(1-t)\big[I(X;Y_1) - \alpha I(X;Z)\big]$$
$$+ t\big[I(X;Y_2) - (1-\alpha)I(X;Z)\big]\Big]$$
$$= \min_{0 \le t \le 1} \max_{0 \le \alpha \le 1} \Big[(1-t)\big[I(X;Y_1) - \alpha I(X;Z)\big]$$
$$+ t\big[I(X;Y_2) - (1-\alpha)I(X;Z)\big]\Big] \quad (199)$$

where the equality follows from the minimax theorem. Now eliminating α in (199) yields for $t \le \frac{1}{2}$ and $t \ge \frac{1}{2}$ the corresponding expressions in (198) which are then equivalent to the original formulation (31). ∎

REFERENCES

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

[2] 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, Jun. 2015, pp. 1971–1975.

[3] C. E. Shannon, "Communication theory of secrecy systems," *Bell Labs 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, no. 8, pp. 1355–1387, 1975.

[5] 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.

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

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

[8] H. V. Poor and R. F. Schaefer, "Wireless physical layer security," *Proc. Nat. Acad. Sci. USA*, vol. 114, no. 1, pp. 19–26, Jan. 2017.

[9] R. F. Schaefer, H. Boche, A. Khisti, and H. V. Poor, Eds., *Information Theoretic Security and Privacy of Information Systems*. Cambridge, U.K.: Cambridge Univ. Press, 2017.

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

[11] E. Ekrem and S. Ulukus, "Capacity region of Gaussian MIMO broadcast channels with common and confidential messages," *IEEE Trans. Inf. Theory*, vol. 58, no. 9, pp. 5669–5680, Sep. 2012.

[12] R. Liu, T. Liu, H. V. Poor, and S. Shamai (Shitz), "New results on multiple-input multiple-output broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. 59, no. 3, pp. 1346–1359, Mar. 2013.

[13] R. F. Wyrembelski, M. Wiese, and H. Boche, "Strong secrecy in bidirectional broadcast channels with confidential messages," *IEEE Trans. Inf. Forensics Security*, vol. 8, no. 2, pp. 324–334, Feb. 2013.

[14] Y. Liang and H. V. Poor, "Multiple-access channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. 54, no. 3, pp. 976–1002, Mar. 2008.

[15] M. Wiese and H. Boche, "Strong secrecy for multiple access channels," in *Information Theory, Combinatorics, and Search Theory*. Berlin, Germany: Springer, 2013, pp. 71–122.

[16] R. Liu, I. Marić, P. Spasojević, and R. D. Yates, "Discrete memoryless interference and broadcast channels with confidential messages: Secrecy rate regions," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2493–2507, Jun. 2008.

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

[18] 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.

[19] W. Kang and N. Liu, "Wiretap channel with shared key," in *Proc. IEEE Inf. Theory Workshop*, Dublin, Ireland, Aug./Sep. 2010, pp. 1–5.

[20] R. Ahlswede and N. Cai, "Transmission, identification and common randomness capacities for wire-tap channels with secure feedback from the decoder," in *General Theory of Information Transfer and Combinatorics* (Lecture Notes in Computer Science), vol. 4123. Berlin, Germany: Springer, 2006, pp. 258–275.

[21] D. Gündüz, D. R. Brown, III, and H. V. Poor, "Secret communication with feedback," in *Proc. Int. Symp. Inf. Theory Appl.*, Auckland, New Zealand, Dec. 2008, pp. 1–6.

[22] E. Ardestanizadeh, M. Franceschetti, T. Javidi, and Y.-H. Kim, "Wiretap channel with secure rate-limited feedback," *IEEE Trans. Inf. Theory*, vol. 55, no. 12, pp. 5353–5361, Dec. 2009.

[23] I. Csiszár, "Almost independence and secrecy capacity," *Problemy Peredachi Inform.*, vol. 32, no. 1, pp. 48–57, 1996.

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

[25] J. L. Massey, "A simplified treatment of Wyner's wire-tap channel," in *Proc. 21st Annu. Allerton Conf. Commun. Control Comput.*, Monticello, IL, USA, Oct. 1983, pp. 268–276.

[26] I. Bjelaković, H. Boche, and J. Sommerfeld, "Capacity results for compound wiretap channels," in *Proc. IEEE Inf. Theory Workshop*, Paraty, Brazil, Oct. 2011, pp. 60–64.

[27] 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.

[28] 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.

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

[30] C. Nair and A. El Gamal, "An outer bound to the capacity region of the broadcast channel," *IEEE Trans. Inf. Theory*, vol. 53, no. 1, pp. 350–355, Jan. 2007.

[31] K. Marton, "A coding theorem for the discrete memoryless broadcast channel," *IEEE Trans. Inf. Theory*, vol. IT-25, no. 3, pp. 306–311, May 1979.

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

[33] S. I. Gelfand and M. S. Pinsker, "Coding for channel with random parameters," *Problems Control Inf. Theory*, vol. 9, no. 1, pp. 19–31, 1980.

[34] G. Kramer, "Topics in multi-user information theory," *Found. Trends Commun. Inf. Theory*, vol. 4, nos. 4–5, pp. 265–444, 2007.

[35] C. Nair and V. W. Zizhou, "On the inner and outer bounds for 2-receiver discrete memoryless broadcast channels," in *Proc. Inf. Theory Appl. Workshop*, San Diego, CA, USA, Jan. 2008, pp. 226–229.

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

[37] A. S. Mansour, R. F. Schaefer, and H. Boche, "Secrecy measures for broadcast channels with receiver side information: Joint vs individual," in *Proc. IEEE Inf. Theory Workshop*, Hobart, TAS, Australia, Nov. 2014, pp. 426–430.

Rafael F. Schaefer (S'08–M'12–SM'17) received the Dipl.Ing. degree in electrical engineering and computer science from the Technische Universität Berlin, Germany, in 2007, and the Dr.-Ing. degree in electrical engineering from the Technische Universität München, Germany, in 2012. From 2007 to 2010, he was a Research and Teaching Assistant with the Technische Universität Berlin and from 2010 to 2013, with the Technische Universität München. From 2013 to 2015, he was a Post-Doctoral Research Fellow with Princeton University. Since 2015, he has been an Assistant Professor with the Technische Universität Berlin. Among his publications is the recent book *Information Theoretic Security and Privacy of Information Systems* (Cambridge University Press, 2017). He was a recipient of the VDE Johann-Philipp-Reis Prize in 2013. He received the best paper award of the German Information Technology Society in 2016. He was one of the exemplary reviewers of the IEEE COMMUNICATION LETTERS in 2013. He is an Associate Member of the IEEE Information Forensics and Security Technical Committee. He is currently an Associate Editor of the IEEE TRANSACTIONS OF COMMUNICATIONS.

Ashish Khisti (S'01–M'09) received the B.A.Sc. degree from the Engineering Sciences Program (Electrical Engineering Option) the University of Toronto, Toronto, ON, Canada, in 2002, and the M.S. and Ph.D. degrees in electrical engineering and computer science from the Massachusetts Institute of Technology, Cambridge, MA, USA, in 2004 and 2009, respectively. He is currently an Associate Professor with the University of Toronto and holds a Canada Research Chair in wireless networks. His research interests include information theory, physical layer security, and error control coding for multimedia applications. He also actively consults telecommunication companies. He was a recipient of the HP-IRP Award from Hewlett-Packard and an Ontario Early Researcher Award. He currently serves as an Associate Editor in Shannon theory for the IEEE TRANSACTIONS ON INFORMATION THEORY.

H. Vincent Poor (S'72–M'77–SM'82–F'87) received the Ph.D. degree in EECS from Princeton University in 1977. From 1977 to 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 currently the Michael Henry Strater University Professor of Electrical Engineering. From 2006 to 2016, he served as the Dean of the Princeton's School of Engineering and Applied Science. He has also held visiting appointments at several other universities, including most recently at Berkeley and Cambridge. His research interests are in the areas of information theory and signal processing, and their applications in wireless networks and related fields, such as energy systems and social networks. Among his publications in these areas is the recent book *Information Theoretic Security and Privacy of Information Systems* (Cambridge University Press, 2017).

Dr. Poor is a Member of the National Academy of Engineering and the National Academy of Sciences and a Foreign Member of the Royal Society. He is also a Fellow of the American Academy of Arts and Sciences, the National Academy of Inventors, and other national and international academies. He received the Marconi Award and the Armstrong Award of the IEEE Communications Society in 2007 and 2009, respectively. Recent recognition of his work includes the 2017 IEEE Alexander Graham Bell Medal, Honorary Professorships at Peking University and Tsinghua University, both conferred in 2017, and the D.Sc. *honoris causa* from Syracuse University in 2017.