A multiple access approach for the compound wiretap channel

Etienne Perron, Suhas Diggavi, Emre Telatar
EPFL, Lausanne, Switzerland
Email: {etienne.perron,suhas.diggavi,emre.telatar}@epfl.ch

***Abstract*—The compound wiretap channel generalizes the classical problem in broadcast information-theoretic secrecy by allowing a class of potential eavesdroppers. This represents uncrtainty in the eavesdropper channel and the characterization of its secrecy capacity is an open question. In this paper we present a new coding scheme that generalizes known approaches to this problem. The scheme prefixes an artificial multiple access channel to the transmission scheme in order to design a structured transmit codebook. The idea is that such a structure can potentially increase the perfect secrecy rate for the legitimate users in the presence of the class of eavesdroppers. We develop the achievable secrecy rates of this scheme and provide examples where this scheme is optimal.**

I. Introduction

Information-theoretic secrecy is predicated on the fact that different versions of the transmitted signal are received by a legitimate receiver and an eavesdropper. It fundamentally exploits this distinction to create secrecy, as shown by Wyner in his seminal paper on wiretap channels [7]. However, it critically uses the knowledge of the channel from the transmitter to the eavesdropper to create secrecy. A natural question posed in [3] is how to generate secrecy when there is uncertainty in the knowledge of the eavesdropper channel. This can be modeled by allowing the eavesdropper channel to belong to a class, and therefore this motivates the terminology, compound wiretap channel.

The secrecy capacity of the compound wiretap channel is still open, though some important special cases have been resolved in [3], [4]. A coding scheme similar to that proposed in [1], is analyzed for the compound wiretap channel in [3]. This scheme was shown to be optimal for the case where all the eavesdroppers are degraded with respect to the legitimate receiver [3]. An interesting new coding scheme was shown to be optimal in [4] for the parallel Gaussian channel. This scheme developed a structured product code, where the overall codebook was chosen as a direct product of sub-codebooks designed for each parallel channel. It was shown that this scheme out-performed the scheme proposed in [3] for Gaussian codebooks. In this paper we explore a scheme that has both the aforementioned coding schemes as special cases. The basic idea is to prefix an artificial multiple access channel at the transmitter, thereby yielding a structured transmission codebook. Another way to interpret this idea is that we introduce multiple auxiliary random variables, instead of the single auxiliary random variable used in [3]. We encapsulate structure on these auxiliary random variables by thinking of them as forming a multiple-access channel to the transmitted symbol. The scheme in [3], [1] can be interpreted as prefixing a point-to-point channel between the auxiliary random variable and the transmitted symbol, making it a special case of this scheme. Clearly the product codebook for parallel channels introduced in [4] is a very special multiple access channel (MAC) which has orthogonal components corresponding to each parallel channel transmit symbol.

Additionally, we allow for the secret message to potentially depend only on a subset of the auxiliary random variables. In this sense, we can interpret the rest of the auxiliary variables as “noise insertion”, aiming to increase the secrecy rate. This idea is similar to one developed in [5] for deterministic relay networks. The main result in this paper is to provide an achievable secrecy rate for the aforementioned scheme when there is a set of legitimate users to whom a common secret message is to be delivered, while keeping it information-theoretically secure from a class of eavesdroppers. The secrecy rate obtained can be optimized over the choice of the prefixed MAC channel, as well as the partitioning between the “noise” and “message” auxiliary variables. Though the achievability is general, we illustrate its application to specific examples.

The paper is organized as follows. We formulate the problem and develop the notation in Section II. We state the main result of the paper in Section III. The proof is given in Section IV. Application of the main result is illustrated through examples in Section V. We conclude with a brief discussion in Section VI.

II. Problem Setup

We consider a discrete memoryless wiretap channel with one transmitter, a class of destinations and a class of possible eavesdroppers. In this section, we define this setup more precisely.

Definition 1: A *discrete compound wiretap channel* is defined by a finite transmit alphabet $\mathcal{X}$, J finite receive alphabets $\mathcal{Y}_1, \ldots, \mathcal{Y}_J$ for the J destinations (the legitimate receivers), K receive alphabets $\mathcal{Z}_1, \ldots, \mathcal{Z}_K$, for the K eavesdroppers, as well as a transmit probability mass function $p_{\{Y_j\}_{j=1}^J, \{Z_k\}_{k=1}^K | X}$. We denote the entities of this channel by S for the transmitter (the source), D_j for the destination j, and E_k for eavesdropper k.

Definition 2: A (T, ϵ)-*code* consists of a possibly random encoding function $f : \mathcal{M} \to \mathcal{X}^T, M \mapsto \mathbf{X}$ and J decoding functions $g_j : \mathcal{Y}_j^T \to \mathcal{M}, \mathbf{Y}_j \mapsto \hat{M}_j$, for $j = 1, \ldots, J$. The message M is uniformly distributed in the message alphabet $\mathcal{M}$ and $\hat{M}_j$ is the estimate of M produced at D_j. We require that the code is ϵ-reliable in the sense that for all $j = 1, \ldots, J$, $\mathbf{P}(\hat{M}_j \neq M) \leq \epsilon$.

Definition 3: The *rate* of a (T, ϵ)-code is defined as $\frac{1}{T} H(M) = \frac{1}{T} \log |\mathcal{M}|$, where the log is base 2, *i.e.*, the rate is measured in bits per channel use.

Definition 4: The *equivocation* of a (T, ϵ)-code is defined as $\frac{1}{T} \min_{k=1,\ldots,K} H(M|\mathbf{Z}_k)$, where $\mathbf{Z}_k$ is the random received sequence at E_k.

Definition 5: A pair of positive numbers (R, R_e) is said to be *achievable* if $R \geq R_e$ and for every $\epsilon > 0$, there exists a (T, ϵ)-code of rate at least $R - \epsilon$ and equivocation at least $R_e - \epsilon$.

Definition 6: The *perfect secrecy capacity* C_s of a compound wiretap channel is the largest number R such that (R, R) is achievable.

III. Main Results

In this section, we state our achievability for the general discrete memoryless compound wiretap channel. We prefix the channel by an artificial multiple-access channel (MAC) at S, which is defined as follows:

Definition 7: An auxiliary *multiple-access channel (MAC)*, denoted by $\mathfrak{C}$ is defined by L discrete alphabets $\mathcal{U}_l$, $l = 1, \ldots, L$ as well as a MAC probability mass function $p_{X|\{U_l\}, l=1,\ldots,L}$.

Instead of encoding M directly into a sequence $\mathbf{X}$, S splits M into L sub-messages M_l, of respective rates R_l, $l = 1, \ldots, L$. These messages are then encoded using a MAC-codebook for the MAC $\mathfrak{C}$ in the same way as the product codebook used in [4]. Some of these messages are then binned using a number of bins. This corresponds to a randomization to confuse the eavesdroppers. The binning is done jointly for all the involved messages of the MAC, again similarly to [4]. Note that the secret message does not correspond to any of the messages M_l but to the bin index. We explain this later in more detail.

The prefixed MAC gives the additional freedom of using a product codebook. By optimizing over all possible MACs, we can hence potentially enlarge the secrecy rate as compared to the rate given in [2]. In addition to the product codebook, we also use the technique of noise insertion. By this, we mean that we generate some messages that are pure noise, and we use a subset of the MAC "users" to transmit these messages. We do not require any destination or any of the eavesdroppers to decode these noise messages. The purpose of the noise messages is to alter the achievable rate-regions for the eavesdroppers. Of course, the rate-region for the destinations may also be altered. The aim here is to carefully choose the "users" over which the noise messages are transmitted, in order to appropriately balance the decrease of total information that the eavesdroppers can gather, with the decrease at the destinations, in order to maximize secrecy rate.

Definition 8: We denote by $\mathcal{L} = \{1, \ldots, L\}$ the set of all users. By $\mathcal{L}_m \subseteq \mathcal{L}$, we denote the set of users that are used to transmit messages. All the nodes in $\mathcal{L}_m^c \triangleq \mathcal{L} \setminus \mathcal{L}_m$ are noise users. Hence, the messages $M_{\mathcal{L}_m^c} = (M_l)_{l \in \mathcal{L}_m^c}$ are not required to be decoded at the destination.

Before stating the achievable secrecy rate, we define the MAC rate region for each destination D_j, $j \in \{1, \ldots, J\}$ and for each eavesdropper E_k, $k \in \{1, \ldots, K\}$ as follows.

Definition 9: For a given $j \in \{1, \ldots, J\}$, a given auxiliary MAC $\mathfrak{C}$, a given product distribution $\prod_{l=1}^{L} p_{U_l}$, a set of message users $\mathcal{L}_m$, and a fixed tuple of noise rates $R_{\mathcal{L}_m^c}$, we define $\tilde{\mathcal{R}}_{D_j}(p, \mathfrak{C}, \mathcal{L}_m, R_{\mathcal{L}_m^c})$ as

$$\tilde{\mathcal{R}}_{D_j}(p, \mathfrak{C}, \mathcal{L}_m, R_{\mathcal{L}_m^c}) = \big\{ R_{\mathcal{L}_m} : \sum_{l \in \mathcal{A}} R_l \leq I(U_{\mathcal{A}}; Y_j | U_{\mathcal{A}^c}) \\ \forall \mathcal{A} \subseteq \mathcal{L} \text{ for which } \mathcal{L}_m \cap \mathcal{A} \neq \phi \big\}.$$

In other words, $\cup_{\prod_l p_{U_l}} \tilde{\mathcal{R}}_{D_j}(p, \mathfrak{C}, \mathcal{L}_m, R_{\mathcal{L}_m^c})$ is the capacity region for the MAC from $U_{\mathcal{L}_m}$ to Y_j if $R_{\mathcal{L}_m^c}$ has already been fixed.

Note that the region described in Definition 9 is only defined by constraints that involve at least one of the message-users. Constraints that are exclusively on the rates of noise-users are not relevant, because the noise rates have already been fixed.

We define a similar region for each eavesdropper:

Definition 10: For a given $k \in \{1, \ldots, K\}$, a given auxiliary MAC $\mathfrak{C}$, a given product distribution $\prod_{l=1}^{L} p_{U_l}$, a set of message users $\mathcal{L}_m$, and a fixed tuple of noise rates $R_{\mathcal{L}_m^c}$, we define $\tilde{\mathcal{R}}_{E_k}(p, \mathfrak{C}, \mathcal{L}_m, R_{\mathcal{L}_m^c})$ as

$$\tilde{\mathcal{R}}_{E_k}(p, \mathfrak{C}, \mathcal{L}_m, R_{\mathcal{L}_m^c}) = \big\{ R_{\mathcal{L}_m} : \sum_{l \in \mathcal{A}} R_l \leq I(U_{\mathcal{A}}; Z_k | U_{\mathcal{A}^c}) \\ \forall \mathcal{A} \subseteq \mathcal{L} \text{ for which } \mathcal{L}_m \cap \mathcal{A} \neq \phi \big\}.$$

In addition, we define the MAC region for the noise rates at each eavesdropper as follows.

Definition 11: For a given $k \in \{1, \ldots, K\}$, a given auxiliary MAC $\mathfrak{C}$, a given product distribution $\prod_{l=1}^{L} p_{U_l}$, and a set of message users $\mathcal{L}_m$, we define $\mathcal{R}_{E_k}(p, \mathfrak{C}, \mathcal{L}_m)$ as

$$\mathcal{R}_{E_k}(p, \mathfrak{C}, \mathcal{L}_m) = \big\{ R_{\mathcal{L}_m^c} : \sum_{l \in \mathcal{A}} R_l \leq I(U_{\mathcal{A}}; Z_k | U_{\mathcal{A}^c}) \\ \forall \mathcal{A} \subseteq \mathcal{L}_m^c \big\}.$$

If a tuple of noise rates $R_{\mathcal{L}_m^c}$ lies in $\mathcal{R}_{E_k}(p, \mathfrak{C}, \mathcal{L}_m)$ as defined here, then it is achievable provided that $R_l = 0$ for all message users $l \in \mathcal{L}_m$.

Note that a rate tuple $R_{\mathcal{L}}$ (noise and message rates combined) lies in the achievable MAC rate region for some eavesdropper E_k if for the given $\mathcal{L}_m$, $R_{\mathcal{L}_m^c}$ satisfies Definition 11 and $R_{\mathcal{L}_m}$ satisfies Definition 10 with $R_{\mathcal{L}_m^c}$ as an argument.

Now we can state the general achievability theorem.

Theorem 1:

$$C_s \geq \max_{\mathfrak{C}} \max_{\prod_l p_{U_l}} \max_{\mathcal{L}_m \subseteq \mathcal{L}} \\ \max_{R_{\mathcal{L}_m^c} \in \cap_k \mathcal{R}_{E_k}(p, \mathfrak{C}, \mathcal{L}_m)} \max_{R_{\mathcal{L}_m} \in \cap_j \tilde{\mathcal{R}}_{D_j}(p, \mathfrak{C}, \mathcal{L}_m, R_{\mathcal{L}_m^c})} \\ \min_{k=1,\ldots,K} \min_{R_{\mathcal{L}_m}^{(k)} \in \tilde{\mathcal{R}}_{E_k}(p, \mathfrak{C}, \mathcal{L}_m, R_{\mathcal{L}_m^c})} \sum_{l \in \mathcal{L}_m} (R_l - R_l^{(k)})^+. \quad (1)$$

IV. Proof of Theorem 1

Fix any MAC $\mathfrak{C}$, a product distribution p of the random variables $U_{\mathcal{L}}$, a set of message users $\mathcal{L}_m$, a tuple of noise rates $R_{\mathcal{L}_m^c}$ in $\cap_k \mathcal{R}_{E_k}(p, \mathfrak{C}, \mathcal{L}_m)$ and a tuple of message rates $R_{\mathcal{L}_m}$ in $\cap_j \tilde{\mathcal{R}}_{D_j}(p, \mathfrak{C}, \mathcal{L}_m, R_{\mathcal{L}_m^c})$. We show that for any such choice, the rate

$$A \triangleq \min_{k=1,\ldots,K} \min_{R^{(k)}_{\mathcal{L}_m} \in \tilde{\mathcal{R}}_{E_k}(p,\mathfrak{C},\mathcal{L}_m,R_{\mathcal{L}_m^c})} \sum_{l \in \mathcal{L}_m} (R_l - R_l^{(k)})^+ \quad (2)$$

is achievable with perfect secrecy.

Code Generation: We generate a random code in the following way. For every user $l \in \mathcal{L}$, we generate a codebook $\mathcal{C}_l$ of rate R_l from $p^T_{U_l}$. The overall codebook $\mathcal{C}_{\mathcal{L}}$ is the direct product of all the codebooks $\mathcal{C}_l$. Also, for the set of message users, we have a codebook $\mathcal{C}_{\mathcal{L}_m} \subseteq \mathcal{C}_{\mathcal{L}}$ that is the direct product of all the $\mathcal{C}_l$, for $l \in \mathcal{L}_m$, and likewise for the noise users $\mathcal{C}_{\mathcal{L}_m^c}$. We now bin the codewords in $\mathcal{C}_{\mathcal{L}_m}$ into 2^{TA} bins.

Encoding: Note that the message will not be mapped to any of the messages $M_{\mathcal{L}}$. Instead, each realization of the message corresponds to one of the bins. Let the message be denoted by W. Once this bin is identified, we pick one codeword from that bin uniformly at random. This codeword is a member of $\mathcal{C}_{\mathcal{L}_m}$. In addition, we pick uniformly at random a codeword from the noise codebook $\mathcal{C}_{\mathcal{L}_m^c}$. The two codewords together are in $\mathcal{C}_{\mathcal{L}}$ and hence are a tuple of L user-codewords. Let $\mathbf{U}_l$ be the user-codeword for user l. At time t, we transmit $(U_1[t], \ldots, U_L[t])$ over the artificial MAC to obtain $X[t]$, which is in turn transmitted over the wiretap channel.

Decoding: Note that $R_{\mathcal{L}_m}$ lies in $\cap_j \tilde{\mathcal{R}}_{D_j}(p, \mathfrak{C}, \mathcal{L}_m, R_{\mathcal{L}_m^c})$. Thus, from the standard error analysis of a randomly generated code, we find that there exists at least one code that allows reliable decoding of $\mathbf{U}_{\mathcal{L}_m}$ at every decoder D_j. Hence, all the D_j are highly likely to identify the correct bin, and hence the correct secret message.

Secrecy Analysis: Without loss of generality, assume that E_1 is the strongest eavesdropper in the sense that $k = 1$ is the minimizer of (2). We first analyze the equivocation at E_1. The same proof can then be used for the weaker eavesdroppers by enhancement of their channel output. Let W be the secret message, *i.e.*, W denotes a bin. The equivocation at E_1 can be lower bounded as

$$\begin{aligned} H(W|\mathbf{Z}_1) &= H(W, \mathbf{U}_{\mathcal{L}}|\mathbf{Z}_1) - H(\mathbf{U}_{\mathcal{L}}|W, \mathbf{Z}_1) \\ &\geq H(\mathbf{U}_{\mathcal{L}}|\mathbf{Z}_1) - H(\mathbf{U}_{\mathcal{L}}|W, \mathbf{Z}_1). \end{aligned} \quad (3)$$

Let us first find an upper bound on the second term in (3). The binning divides each sub-codebook $\mathcal{C}_l$ into 2^{TA} parts. Assume that the bin index W is known at E_1. In this case, E_1 can recover the transmit messages $M_{\mathcal{L}}$ (and hence the transmitted codewords $\mathbf{U}_{\mathcal{L}}$) with high probability (over all codes) if $\forall \mathcal{A} \subseteq \mathcal{L}$. We will use this insight to show that the following holds.

$$\sum_{l \in \mathcal{A}} R_l - A \mathbf{1}_{\{\mathcal{L}_m \cap \mathcal{A} \neq \phi\}} \leq I(U_{\mathcal{A}}; Z_1 | U_{\mathcal{A}^c}). \quad (4)$$

The indicator function in (4) is present only for subsets $\mathcal{A}$ that concern codewords that have been binned. Since $R_{\mathcal{L}_m^c}$ was chosen from $\cap_k \mathcal{R}_{E_k}(p, \mathfrak{C}, \mathcal{L}_m)$, (4) holds for all $\mathcal{A}$ such that $\mathcal{A} \cap \mathcal{L}_m = \phi$. Now, we show that (4) holds. Let $R^*_{\mathcal{L}_m} \in \tilde{\mathcal{R}}_{E_1}(p, \mathfrak{C}, \mathcal{L}_m, R_{\mathcal{L}_m^c})$ be one of the minimizers of the inner minimization in (2) when $k = 1$, *i.e.*, we have

$$A = \sum_{l \in \mathcal{L}_m} (R_l - R_l^*)^+.$$

Now, fix any $\mathcal{A} \subseteq \mathcal{L}$ such that $\mathcal{A} \cap \mathcal{L}_m \neq \phi$. We have

$$\begin{aligned} A &\geq \sum_{l \in \mathcal{A} \cap \mathcal{L}_m} (R_l - R_l^*)^+ \\ &\geq \sum_{l \in \mathcal{A} \cap \mathcal{L}_m} (R_l - R_l^*). \end{aligned}$$

Using this, we obtain

$$\begin{aligned} \sum_{l \in \mathcal{A}} R_l - A &= \sum_{l \in \mathcal{A} \cap \mathcal{L}_m} R_l - A + \sum_{l \in \mathcal{A} \setminus \mathcal{L}_m} R_l \\ &\leq \sum_{l \in \mathcal{A} \cap \mathcal{L}_m} R_l^* + \sum_{l \in \mathcal{A} \setminus \mathcal{L}_m} R_l \\ &\leq I(U_{\mathcal{A}}; Z_1 | U_{\mathcal{A}^c}), \end{aligned}$$

where the last inequality is true because of the following. We know that $R^*_{\mathcal{L}_m} \in \tilde{\mathcal{R}}_{E_1}(p, \mathfrak{C}, \mathcal{L}_m, R_{\mathcal{L}_m^c})$, where the argument $R_{\mathcal{L}_m^c}$ is the same tuple of noise messages that is used in the above expression. Hence, the inequality follows from Definition 10. Hence, (4) is satisfied for all choices of $\mathcal{A}$, and hence, there exists a code under which E_1 can reliably decode $M_{\mathcal{L}}$. From Fano's inequality, it follows that

$$H(\mathbf{U}_{\mathcal{L}}|W, \mathbf{Z}_1) \leq T\epsilon_1. \quad (5)$$

Let us now lower bound the first term in (3). Let $R^*_{\mathcal{L}_m}$ be as defined before, *i.e.*, we have $A = \sum_{l \in \mathcal{L}_m} (R_l - R_l^*)^+$. We know that $R^*_{\mathcal{L}_m} \in \tilde{\mathcal{R}}_{E_1}(p, \mathfrak{C}, \mathcal{L}_m, R_{\mathcal{L}_m^c})$. It is well known [6] that $\tilde{\mathcal{R}}_{E_1}(p, \mathfrak{C}, \mathcal{L}_m, R_{\mathcal{L}_m^c})$ is a polymatroid. Hence, any $R'_{\mathcal{L}_m}$ such that $R'_l \leq R^*_l$ for all $l \in \mathcal{L}_m$ is also in $\tilde{\mathcal{R}}_{E_1}(p, \mathfrak{C}, \mathcal{L}_m, R_{\mathcal{L}_m^c})$. In particular, if we define

$$R'_l \triangleq \begin{cases} R^*_l & \text{if } R^*_l \leq R_l \\ R_l & \text{otherwise,} \end{cases}$$

then this choice of $R'_{\mathcal{L}_m}$ lies in $\tilde{\mathcal{R}}_{E_1}(p, \mathfrak{C}, \mathcal{L}_m, R_{\mathcal{L}_m^c})$. Define

$$\mathcal{L}'_m \triangleq \{l \in \mathcal{L}_m : R^*_l < R_l\}.$$

Note that by construction of $R'_{\mathcal{L}_m}$, we have

$$\begin{aligned} A &= \sum_{l \in \mathcal{L}_m} (R_l - R'_l) \\ &= \sum_{l \in \mathcal{L}'_m} (R_l - R'_l), \end{aligned} \quad (6)$$

where the first identity is true because R'_l is never larger than R_l, and in the second identity, we dropped some terms that are zero. Let $\mathcal{A}' \subseteq \mathcal{L}$ be such that

1)

$$\sum_{l \in \mathcal{A}' \cap \mathcal{L}_m} R'_l + \sum_{l \in \mathcal{A}' \cap \mathcal{L}_m^c} R_l = I(U_{\mathcal{A}'}; Z_1 | U_{\mathcal{A}'}) \quad (7)$$

is satisfied with equality,
2) $\mathcal{L}'_m \subseteq \mathcal{A}'$ and none of the nodes in $\mathcal{L}_m \setminus \mathcal{L}'_m$ is in $\mathcal{A}'$,
3) and $\mathcal{A}'$ might contain a certain number of noise users (users from $\mathcal{L}^c_m$).

Such a set $\mathcal{A}'$ exists for the following reason. Since for any $l \in \mathcal{L}'_m$, $R^*_l \leq R_l$, we could potentially make (2) smaller by increasing R^*_l. However, we know by definition that $R^*_{\mathcal{L}_m}$ is an optimizer of (2). Hence, one of the constraints in $\tilde{\mathcal{R}}_{E_1}(p, \mathfrak{C}, \mathcal{L}_m, R_{\mathcal{L}^c_m})$ must be active in the l-direction. On the other hand, for any $l \in \mathcal{L}_m \setminus \mathcal{L}'_m$, we obtained R'_l by decreasing R^*_l, which was at most at the boundary of the region $\tilde{\mathcal{R}}_{E_1}(p, \mathfrak{C}, \mathcal{L}_m, R_{\mathcal{L}^c_m})$. Hence, R'_l is certainly away from that boundary, and none of the constraints in $\tilde{\mathcal{R}}_{E_1}(p, \mathfrak{C}, \mathcal{L}_m, R_{\mathcal{L}^c_m})$ is active in the l direction. Since these two facts hold for all $l \in \mathcal{L}_m$, it follows that there exists a dominating hyperplane in the boundary of $\tilde{\mathcal{R}}_{E_1}(p, \mathfrak{C}, \mathcal{L}_m, R_{\mathcal{L}^c_m})$ that acts only on $R_{\mathcal{L}'_m}$ and maybe on some of the noise rates in $R_{\mathcal{L}^c_m}$. The set $\mathcal{A}'$ is simply such that (7) corresponds to a description of this hyperplane. Now, we can lower bound the first term in (3):

$$\begin{aligned} H(\mathbf{U}_{\mathcal{L}}|\mathbf{Z}_1) &= H(\mathbf{U}_{\mathcal{A}'^c}|\mathbf{Z}_1) + H(\mathbf{U}_{\mathcal{A}'}|\mathbf{Z}_1, \mathbf{U}_{\mathcal{A}'^c}) \\ &\geq H(\mathbf{U}_{\mathcal{A}'}|\mathbf{Z}_1, \mathbf{U}_{\mathcal{A}'^c}) \\ &= H(\mathbf{U}_{\mathcal{A}'}|\mathbf{U}_{\mathcal{A}'^c}) - I(\mathbf{U}_{\mathcal{A}'}; \mathbf{Z}_1|\mathbf{U}_{\mathcal{A}'^c}) \\ &\overset{(a)}{=} H(\mathbf{U}_{\mathcal{A}'}) - I(\mathbf{U}_{\mathcal{A}'}; \mathbf{Z}_1|\mathbf{U}_{\mathcal{A}'^c}) \\ &\overset{(b)}{\geq} T \sum_{l \in \mathcal{A}'} R_l - T I(U_{\mathcal{A}'}; Z_1|U_{\mathcal{A}'^c}) \\ &\overset{(c)}{=} T \sum_{l \in \mathcal{L}'_m} R_l - T\Big(I(U_{\mathcal{A}'}; Z_1|U_{\mathcal{A}'^c}) - \sum_{l \in \mathcal{A}' \setminus \mathcal{L}_m} R_l\Big) \\ &\overset{(d)}{=} T \sum_{l \in \mathcal{L}'_m} R_l - T \sum_{l \in \mathcal{A}' \cap \mathcal{L}'_m} R'_l \\ &\overset{(c)}{=} T \sum_{l \in \mathcal{L}'_m} (R_l - R'_l) \\ &\overset{(e)}{=} TA. \end{aligned} \tag{8}$$

In the above derivation, (a) follows from the fact that the sub-codewords of a product codebook are independent, (b) is obtained by applying the single letter upper bound for a MAC channel to the second term, (c) is true because $\mathcal{A}' = \mathcal{L}'_m \cup \mathcal{A}' \setminus \mathcal{L}_m$ by definition, (d) follows from (7), and (e) follows from (6). Finally, we replace (5) and (8) in (3) to obtain

$$\frac{1}{T} H(W|\mathbf{Z}_1) \geq A - \epsilon_1,$$

where ϵ_1 can be chosen arbitrarily small. Since A is the rate of the message W, this proves perfect secrecy for eavesdropper E_1.

It is now easy to analyze the equivocation of the weaker eavesdroppers. Assume that E_2 is weaker than E_1 in the sense that

$$\begin{aligned} &\min_{R^{(2)}_{\mathcal{L}_m} \in \tilde{\mathcal{R}}_{E_2}(p, \mathfrak{C}, \mathcal{L}_m, R_{\mathcal{L}^c_m})} \sum_{l \in \mathcal{L}_m} (R_l - R^{(2)}_l)^+ < \\ &\min_{R^{(1)}_{\mathcal{L}_m} \in \tilde{\mathcal{R}}_{E_1}(p, \mathfrak{C}, \mathcal{L}_m, R_{\mathcal{L}^c_m})} \sum_{l \in \mathcal{L}_m} (R_l - R^{(1)}_l)^+. \end{aligned} \tag{9}$$

Following Lemma 2 in [3], we can define an enhanced eavesdropper $\tilde{E}_2$ whose channel output is $(Z_2, \tilde{Z})$, where $\tilde{Z}$ is chosen such that (9) is satisfied with equality if we replace E_2 by $\tilde{E}_2$. Furthermore, $X \,-\!\!\circ\!\!-\, (Z_1, Z_2) \,-\!\!\circ\!\!-\, \tilde{Z}$ forms a Markov chain. Now, $\tilde{E}_2$ is among the strongest eavesdroppers, and we can carry out the same analysis as for E_1 to conclude that

$$H(W|\mathbf{Z}_2, \tilde{\mathbf{Z}}) \geq A - \epsilon_2,$$

where ϵ_2 can be chosen arbitrarily small. But since $H(W|\mathbf{Z}_2) \geq H(W|\mathbf{Z}_2, \tilde{\mathbf{Z}})$, this implies that the equivocation at E_2 can be made arbitrarily close to A. Since A is the rate of W, this concludes the secrecy proof.

V. Examples

A. Example 1

Assume that there is only one destination D and two eavesdroppers E_1 and E_2. The received signal at D is simply denoted by Y. Let $\mathcal{X} = \{0, 1\}^2$ and denote the two bits of X as a two-dimensional column vector $X = (X_1, X_2)^\dagger$. Let $Y = (X_1 + N_1, X_2 + N_2)^\dagger$, where N_1 and N_2 are independent Bernoulli-p random variables, and independent of X. Furthermore, let $Z_1 = X_1$ and $Z_2 = X_2$. This example is a binary parallel compound wiretap channel. This is very similar to the Gaussian example given in [4]. We choose to prefix a MAC with orthogonal components such that $(X_1, X_2)^\dagger = (U_1, U_2)^\dagger$. Hence, we can avoid using U_1 and U_2, which simplifies the notation. We set X_1 and X_2 to be independent Bernoulli-$\frac{1}{2}$ random variables. There are no noise users, *i.e.*, $\mathcal{L}_m = \mathcal{L} = \{1, 2\}$. The regions $\tilde{\mathcal{R}}_i(p, \mathfrak{C}, \{1, 2\}, \phi)$, for $i \in \{D, E_1, E_2\}$, are shown in Figure 1. They are denoted by $\tilde{\mathcal{R}}_D$, $\tilde{\mathcal{R}}_{E_1}$ and $\tilde{\mathcal{R}}_{E_2}$, respectively. The region $\tilde{\mathcal{R}}_D$ is the shaded square, while $\tilde{\mathcal{R}}_{E_k}$ is the interval $[0, 1]$ on the axis of user k, for $k = 1, 2$, respectively. The product scheme achieves a secrecy rate of

$$R_s = \min_{k=1,2} \min_{(R^{(k)}_1, R^{(k)}_2) \in \tilde{\mathcal{R}}_{E_k}} \Big((R_1 - R^{(k)}_1)^+ + (R_2 - R^{(k)}_2)^+\Big),$$

which turns out to be $1 - h(p)$. We can easily show that $1 - h(p)$ is actually the secrecy capacity of this channel, by using Theorem 2 in [4]:

$$C_s \leq \max_{p_X} \min_{k=1,2} I(X; Y|Z_k). \tag{10}$$

To evaluate this upper bound, we first maximize $I(X;Y|Z_1)$ over p_X:

$$\begin{aligned} I(X;Y|Z_1) &= H(Y|X_1) - H(Y|X_1,X_2) \\ &= H(Y|X_1) - H(N_1,N_2) \\ &\overset{(a)}{\le} H(X_1+N_1|X_1) + H(X_2+N_2) - H(N_1) - H(N_2) \\ &= H(X_2+N_2) - H(N_2) \\ &\overset{(b)}{\le} 1 - h(p). \end{aligned}$$

Inequality (a) is satisfied with equality if X_1 and X_2 are independent, and inequality (b) if X_2 is a Bernoulli-$\frac{1}{2}$ variable. By symmetry, $I(X;Y|Z_2)$ is maximized by any p_X such that $X_2 \sim \mathcal{B}(\frac{1}{2})$ and $X_1 \perp X_2$. It follows that the i.i.d. Bernoulli-$\frac{1}{2}$ distribution of (X_1,X_2) maximizes both $I(X;Y|Z_1)$ and $I(X;Y|Z_2)$ simultaneously. Hence, the same distribution is the maximizer of (10), and we have $C_s \le 1 - h(p)$. This proves that our scheme achieves the perfect secrecy capacity of the given channel.

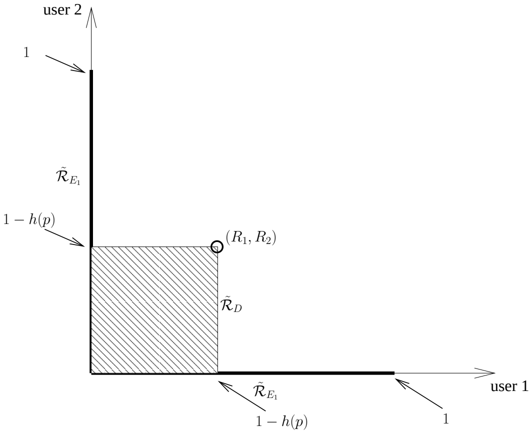

Fig. 1. The MAC-regions used in Example 1.

B. Example 2

Now, consider a deterministic wiretap channel where $X = (X_1,X_2)^\dagger$ takes values in $\{0,1\}^2$. The destination D observes $Y = X_1 + X_2$ and eavesdropper E_k observes $Z_k = X_k$, respectively, for $k = 1,2$. The secrecy capacity for deterministic channels has been found in [3] and turns out to equal 1 bit. Here, we present a scheme that achieves the same secrecy rate and that uses an artificial MAC with noise insertion. One way to prefix a MAC to this channel is to define three user variables U_l, $l = 1,2,3$, and to set up a MAC such that $X = U_1(1,0)^\dagger + U_2(0,1)^\dagger + U_3(1,1)^\dagger$. Note that $(1,1)^\dagger$ lies in the nullspace of the channel transfer matrix of D, *i.e.*, information "modulated" onto $(1,1)^\dagger$ will not affect Y. Hence, user 3 is a good candidate for a noise user. Let p_{U_1,U_2,U_3} be the i.i.d. Bernoulli-$\frac{1}{2}$ distribution. Because the noise does not affect D, we have that for any given noise rate R_3, $\tilde{\mathcal{R}}_D(p,\mathfrak{C},\{1,2\},R_3)$ is the same triangle in the (R_1,R_2)-plane. The volume that one obtains by varying R_3 in $[0,1]$ is depicted in Figure 2 (a cylinder with a triangular base) and denoted by $\tilde{\mathcal{R}}_D$. For eavesdropper E_k, $k = 1,2$, $\tilde{\mathcal{R}}_{E_k}(p,\mathfrak{C},\{1,2\},R_3)$ is a one-dimensional interval on the R_k-axis. As R_3 increases, the length of the interval decreases linearly. This situation is illustrated through the two triangles in Figure 2, denoted by $\tilde{\mathcal{R}}_k$, $k = 1,2$. It is now immediately clear that $R_3 = 1$ is the best choice for the noise rate, because in that case, the MAC regions for E_1 and E_2 reduce to the point 0, while the MAC region for D remains unchanged. The achievable secrecy rate is therefore the largest sum-rate possible in the MAC-region for D, which is 1 bit.

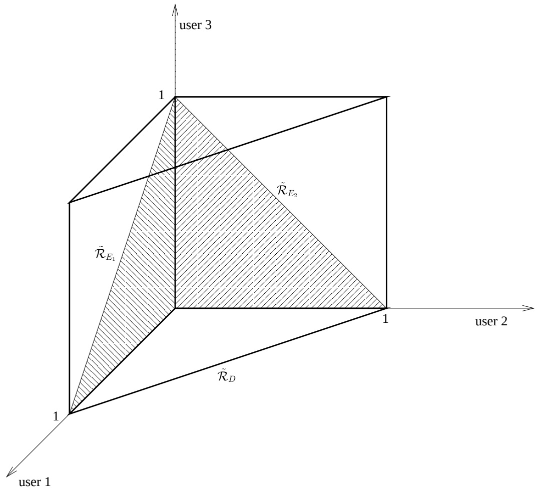

Fig. 2. The MAC-regions used in Example 2.

VI. Discussion

In this paper we introduced a new coding scheme for the compound wiretap channel. The basic idea is to prefix an artificial multiple access channel at the transmitter in order to design a structured code. This scheme contains as special cases all known schemes for the compound wiretap channel. We believe that this is a generalization of the known schemes, and that it will provide an optimal structure for some important special cases. These ideas are part of ongoing work on this problem.

References

[1] I. Csiszár and J. Körner, "Broadcast channels with confidential messages," *IEEE Trans. Inform. Theory*, vol. 24, no. 3, May 1978.

[2] Y. Liang, G. Kramer, H. Poor, and S. Shamai, "Compound wire-tap channels," in *Proc. of the Allerton Conf. on Commun., Control and Computing*, Allerton, USA, 2007.

[3] ——, "Compound wire-tap channels," *EURASIP Journal on Wireless Communications and Networking, Special Issue on Wireless Physical Layer Security*, Dec. 2008, submitted.

[4] T. Liu, V. Prabhakaran, and S. Vishwanath, "The secrecy capacity of a class of parallel gaussian compound wiretap channels," in *Proc. of the IEEE Int. Symposium on Inform. Theory*, Toronto, Canada, Jul. 2008.

[5] E. Perron, S. Diggavi, and E. Telatar, "On noise insertion strategies for wireless network secrecy," in *Proc. of the Information Theory and Applications Workshop*, San Diego, USA, Feb. 2009.

[6] D. Tse and S. Hanly, "Multiaccess fading channels. I. Polymatroid structure, optimal resource allocation and throughput capacities," *IEEE Trans. Inform. Theory*, vol. 44, no. 7, pp. 2796–2815, Nov. 1998.

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