Strong Secrecy in Arbitrarily Varying Wiretap Channels

Igor Bjelaković, Holger Boche, and Jochen Sommerfeld

Lehrstuhl für theoretische Informationstechnik, Technische Universität München, 80290 München, Germany
Email: {igor.bjelakovic, boche, jochen.sommerfeld}@tum.de

***Abstract*— In this work the arbitrarily varying wiretap channel AVWC under the average error criterion and the strong secrecy criterion is studied. We show that in the case of a non-symmetrisable channel to the legitimate receiver the deterministic code secrecy capacity equals the random code secrecy capacity and thus we establish a result for the AVWC similar to that of Ahlswede's dichotomy for ordinary AVCs. We derive a lower bound on the random code secrecy capacity in the case of a best channel to the eavesdropper. We further prove upper bounds on the deterministic code secrecy capacity, which in special cases results in explicit expressions of the secrecy capacity.**

I. Introduction

Models of communication systems connecting the requirement of security against a potential eavesdropper and reliable information transmission to legitimate receivers which suffer from channel uncertainty, have received much interest in current research. Compound wiretap channels were the topic of previous works of the authors [1], [2] and for example of [3], [4]. Here we consider the model of an AVWC where the channel state to both the legitimate receiver and the eavesdropper varies from symbol to symbol in an arbitrary manner, which is unknown to the legitimate participants. Thus apart from eavesdropping the model can simulate an active jamming situation in which the jammer chooses the states. We derive capacity results for the AVWC under the average error probability criterion and a strong secrecy criterion. For the latter we give a operational meaning in terms of the error probability at the eavesdropper.

Our contributions are summarised as follows. We derive a lower bound on the random code secrecy capacity under the strong secrecy criterion in the special case of a "best" channel to the eavesdropper. The proof relies on the coding technique we used for corresponding results for compound wiretap channels in [2]. Then the link to the AVWC is given by the *robustification technique* of Ahlswede [5]. The proof differs significantly from that in [6], where the same capacity result is given under the weak secrecy criterion. In Section III-C we use the *elimination technique* to establish a result for AVWCs similar to Ahlswede's dichotomy [7] which relates the deterministic code secrecy capacity to the random code secrecy capacity. In Section III-D we give both single-letter and multi-letter upper bounds on the deterministic code secrecy capacity, which in special cases leads to a full coding theorem.

II. Arbitrarily Varying Wiretap Channels

A. Definitions

Let A, B, C be finite sets and consider a not necessarily finite family of channels $W_s : A \to \mathcal{P}(B)$ where $s \in S$ denotes the state of the channel. Now, given $s^n = (s_1, s_2, \ldots, s_n) \in S^n$ we define the stochastic matrix

$$W^n(y^n|x^n, s^n) := \prod_{i=1}^{n} W(y_i|x_i, s_i) := \prod_{i=1}^{n} W_{s_i}(y_i|x_i) \quad (1)$$

for all $y^n = (y_1, \ldots, y_n) \in B^n$ and $x^n = (x_1, \ldots, x_n) \in A^n$. An ordinary AVC then is defined as the sequence $\{\mathcal{W}^n\}_{n=1}^{\infty}$ of the family of channels $\mathcal{W}^n = \{W^n(\cdot|\cdot, s^n) : s^n \in S^n\}$. Now let $\mathcal{W}^n$ represent the communication link to a legitimate receiver to which the transmitter wants to send a private message, such that a possible second receiver should be kept as ignorant of that message as possible. We call this receiver the eavesdropper, which observes the output of a second family of channels $\mathcal{V}^n = \{V^n(\cdot|\cdot, s^n) : s^n \in S^n\}$ with the definition of $V^n(\cdot|\cdot, s^n)$ as in (1) for $V_s : A \to \mathcal{P}(C)$, $s \in S$. Then we denote the families of pairs of channels with common input by $\mathfrak{W} = \{(W_{s^n}, V_{s^n}) : s^n \in S^n\}$ and call it the arbitrarily varying wiretap channel. In addition, we assume that the state sequence s^n is unknown to the legitimate receiver, whereas the eavesdropper always knows which channel is in use.

A (n, J_n) code $\mathcal{C}_n$ for the AVWC $\mathfrak{W}$ consists of a stochastic encoder $E : \mathcal{J}_n \to \mathcal{P}(A^n)$ (a stochastic matrix) with a message set $\mathcal{J}_n := \{1, \ldots, J_n\}$ and a collection of mutually disjoint decoding sets $\{D_j \subset B^n : j \in \mathcal{J}_n\}$. The average error probability of a code $\mathcal{C}_n$ is given by

$$e(\mathcal{C}_n) := \max_{s^n \in S^n} \frac{1}{J_n} \sum_{j=1}^{J_n} \sum_{x^n \in A^n} E(x^n|j) W_{s^n}^{\otimes n}(D_j^c|x^n). \quad (2)$$

A *correlated random* (n, J_n, Γ, μ) *code* $\mathcal{C}_n^{\text{ran}}$ for the AVWC is given by a family of wiretap codes $\{\mathcal{C}_n(\gamma)\}_{\gamma \in \Gamma}$ together with a random experiment choosing γ according to a distribution μ on Γ. The mean average error probability of a random code $\mathcal{C}_n^{\text{ran}}$ is defined analogously to the ordinary one but with respect to the random selection by

$$\bar{e}(\mathcal{C}_n^{\text{ran}}) := \max_{s^n \in S^n} \frac{1}{J_n} \sum_{j=1}^{J_n} \sum_{\gamma \in \Gamma} e_j(s^n, \mathcal{C}_n(\gamma))\, \mu(\gamma).$$

Definition 2.1: A non-negative number R is an achievable secrecy rate for the AVWC $\mathfrak{W}$, if there is a sequence $(\mathcal{C}_n)_{n \in \mathbb{N}}$ of (n, J_n) codes such that

$$\liminf_{n \to \infty} \frac{1}{n} \log J_n \geq R \ ,$$

$$\lim_{n \to \infty} e(\mathcal{C}_n) = 0 \quad \text{and,} \quad \lim_{n \to \infty} \max_{s^n \in S^n} I(p_J; V_{s^n}^n) = 0.$$

978-1-4673-0223-4/12/$31.00 ©2012 IEEE

Here J is a uniformly distributed RV taking values in $\mathcal{J}_n$. The secrecy capacity is given as the largest achievable secrecy rate and is denoted by $C_S(\mathfrak{W})$. Analogously we define the secrecy rates and the secrecy capacity for random codes $C_{S,\text{ran}}(\mathfrak{W})$, if we replace $\mathcal{C}_n$ by $\mathcal{C}_n^{\text{ran}}$ in the above definition.
Note that the operational meaning of the strong secrecy is that the average error probability of every decoding strategy of the eavesdropper tends to one in the limit of $J_n \to \infty$, see [2].

III. Capacity Results

A. Preliminaries

For the optimal random coding strategy of the AVWC we need the *robustification technique* by Ahlswede [5]. Therefor let Σ_n be the group of permutations acting on $(1, 2, \ldots, n)$. Then every permutation $\sigma \in \Sigma_n$ induces a bijection $\pi \in \Pi_n$ defined by $\pi : \mathcal{S}^n \to \mathcal{S}^n$ with $\pi(s^n) = (s_{\sigma(1)}, \ldots, s_{\sigma(n)})$ for all $s^n \in \mathcal{S}^n$ and Π_n denotes the group of these bijections.

Lemma 3.1: (Robustification technique) If a function $f : \mathcal{S}^n \to [0, 1]$ satisfies

$$\sum_{s^n \in \mathcal{S}^n} f(s^n) q(s_1) \cdot \ldots \cdot q(s_n) \geq 1 - \gamma \tag{3}$$

for all $q \in \mathcal{P}_0(n, \mathcal{S})$ and some $\gamma \in [0, 1]$, then

$$\frac{1}{n!} \sum_{\pi \in \Pi_n} f(\pi(s^n)) \geq 1 - 3 \cdot (n+1)^{|\mathcal{S}|} \cdot \gamma, \quad \forall s^n \in \mathcal{S}^n \ . \tag{4}$$

Proof: The proof is given in [5]. ∎

To establish the main result for the deterministic code secrecy capacity of the AVWC $\mathfrak{W}$ we need the concept of symmetrisability, which was introduced for AVCs in [8].

Definition 3.2: An AVC is symmetrisable if for some channel $U : A \to S$

$$\sum_{s \in S} W(y|x, s) U(s|x') = \sum_{s \in S} W(y|x', s) U(s|x) \tag{5}$$

for all $x, x' \in A$, $y \in B$.
The authors of [9] proved the following theorem which is a strengthening of Ahlswede's dichotomy result for single-user AVCs, which states that the deterministic code capacity C is either $C = 0$ or equals the random code capacity.

Theorem 3.3: [9] $C > 0$ if and only if the AVC is non-symmetrisable. If $C > 0$, then

$$C = \max_{p \in \mathcal{P}(A)} \min_{W \in \bar{\mathcal{W}}} I(p, W) \tag{6}$$

Here the RHS gives the random code capacity and $\bar{\mathcal{W}}$ denotes the convex closure of all channels.

B. Random Code Construction

First let us define the convex hull of the set of channels $\{W_s : s \in S\}$ by the set of channels $\{W_q : q \in \mathcal{P}(S)\}$, where W_q is defined by

$$W_q(y|x) = \sum_{s \in S} W(y|x, s) q(s), \tag{7}$$

for all possible distributions $q \in \mathcal{P}(S)$. Accordingly we define V_q and its convex hull $\{V_q : q \in \mathcal{P}(S)\}$. Then we denote the convex closure of the set of channels $\{(W_s, V_s) : s \in S\}$ by $\overline{\mathfrak{W}} := \{(W_q, V_q) : q \in \mathcal{P}(\tilde{S}), \tilde{S} \subseteq S, \tilde{S} \text{ is finite}\}$.

Lemma 3.4: The secrecy capacity $C_S(\mathfrak{W})$ of the AVWC $\mathfrak{W}$ equals the secrecy capacity of the AVWC $\overline{\mathfrak{W}}$.
The proof for the AVWC under the strong secrecy criterion can be carried out as in the case of the weak secrecy criterion [10]. Now we can start the construction of the random code of the AVWC $\mathfrak{W}$ under the strong secrecy criterion which differs significantly from the case with the weaker secrecy criterion.

Definition 3.5: We call a channel to the eavesdropper a best channel if there exist a channel $V_{q^*} \in \{V_q : q \in \mathcal{P}(S)\}$ such that all other channels from the set are degraded versions of V_{q^*}. If we denote the output of any channel V_q by Z_q we get

$$X \to Z_{q^*} \to Z_q, \quad \forall q \in \mathcal{P}(S). \tag{8}$$

Proposition 3.6: Provided that there exist a best channel to the eavesdropper, for the random code secrecy capacity $C_{S,\text{ran}}(\mathfrak{W})$ of the AVWC $\mathfrak{W}$ it holds that

$$C_{S,\text{ran}}(\mathfrak{W}) \geq \max_{p \in \mathcal{P}(A)} \big(\min_{q \in \mathcal{P}(S)} I(p, W_q) - \max_{q \in \mathcal{P}(S)} I(p, V_q) \big). \tag{9}$$

Proof: The proof is based on Ahlswedes *robustification technique* [5] and is divided in two parts:
step 1): The set $\overline{\mathcal{W}} := \{(W_q^{\otimes n}, V_q^{\otimes n}) : q \in \mathcal{P}(S)\}$ corresponds to a compound wiretap channel indexed by the set of all possible distributions $q \in \mathcal{P}(S)$. Then in [2] it was shown that for the compound wiretap channel $\overline{\mathcal{W}}$ without channel state information the secrecy capacity is lower bounded by

$$C_{S,\text{comp}}(\overline{\mathcal{W}}) \geq \max_{p \in \mathcal{P}(A)} \big(\min_{q \in \mathcal{P}(S)} I(p, W_q) - \max_{q \in \mathcal{P}(S)} I(p, V_q) \big).$$

More precisely, we define

$$J_n := \lfloor 2^{n[\inf_{q \in \mathcal{P}(S)} I(p, W_q) - \sup_{q \in \mathcal{P}(S)} I(p, V_q) - \tau]} \rfloor \tag{10}$$

$$L_n := \lfloor 2^{n[\sup_{q \in \mathcal{P}(S)} I(p, V_q) + \frac{\tau}{4}]} \rfloor, \tag{11}$$

where $J_n = |\mathcal{J}_n|$ and $l \in [L_n]$ serves as the randomisation parameter, cf. Theorem 3.6 [2]. Then there exist a code $\{(x_{jl}, D_j) : j \in [J_n], l \in [L_n]\}$ with average error probability

$$\frac{1}{J_n} \sum_{j \in [J_n]} \frac{1}{L_n} \sum_{l \in [L_n]} W_q^{\otimes n}(D_j^c | x_{jl}) \leq 2^{-na} \tag{12}$$

for $n \in \mathbb{N}$ sufficiently large and for all $q \in \mathcal{P}(S)$ and $a > 0$, and the secrecy criterion is fulfilled by

$$I(p_J; V_q^{\otimes n}) \leq \epsilon' \tag{13}$$

uniformly in $q \in \mathcal{P}(S)$, where J is a random variable uniformly distributed on the message set $\mathcal{J}_n$ and $\epsilon' > 0$ decreases exponentially in $n \in \mathbb{N}$. Now additionally we assume that there exist a best channel to the eavesdropper V_{q^*} in contrast to the proof in [2]. Hence by the definition of V_{q^*} in (8) and because the mutual Information $I(p, V)$ is convex in V and every member of $\{V_q\}_{q \in \mathcal{P}(S)}$ is a convex combination of the set $\{V_s\}_{s \in S}$, it holds that

$$I(p, V_{q^*}) = \sup_s I(p, V_s) = \sup_{q \in \mathcal{P}(S)} I(p, V_q) \tag{14}$$

for all $p \in \mathcal{P}(A)$. Note that because of (14) for $|S| \leq \infty$ $V_{q^*} \in \{V_s : s \in S\}$. Now we have found that there exist a deterministic code such that for all positive numbers R_S with

$$R_S \leq \inf_{q \in \mathcal{P}(S)} I(p, W_q) - I(p, V_{q^*}) \tag{15}$$

are achievable secrecy rates of the compound wiretap channel. *step 2*): *Robustification* : Now we use the robustification technique to derive from the deterministic code $\mathcal{C}_{\overline{\mathcal{W}}} = \{(x_{jl}, D_j) : j \in [J_n], l \in [L_n]\}$ of the compound wiretap channel $\overline{\mathcal{W}}$ a random code for the AVWC $\mathfrak{W}$, which achieves the same secrecy rates. We note first that by (14) and (13)

$$\max_{s^n \in S^n} I(p_J, V_{s^n}) = I(p_J, V_{q^*}^{\otimes n}) \leq \epsilon', \tag{16}$$

which means, that, due to the assumption of a best channel to the eavesdropper, the code achieving the secrecy rate for the best channel to the eavesdropper fulfills the secrecy criterion for a channel with any state sequence $s^n \in S^n$. Now with (12) it holds that

$$\frac{1}{J_n} \sum_{j \in [J_n]} \frac{1}{L_n} \sum_{l \in [L_n]} \sum_{s^n \in S^n} W^{\otimes n}(D_j | x_{jl}, s^n) q^n(s^n) \geq 1 - 2^{-na}$$

for all $q \in \mathcal{P}_0(n, S)$. Now let $\pi \in \Pi_n$ be the bijection on S^n induced by the permutation $\sigma \in \Sigma_n$. Since (3) is fulfilled with

$$f(s^n) = \frac{1}{J_n} \sum_{j \in [J_n]} \frac{1}{L_n} \sum_{l \in [L_n]} W^{\otimes n}(D_j | x_{jl}, s^n) \tag{17}$$

it follows from (4) that

$$\frac{1}{n!} \sum_{\pi \in \Pi_n} \frac{1}{J_n} \sum_{j \in [J_n]} \frac{1}{L_n} \sum_{l \in [L_n]} W^{\otimes n}(D_j | x_{jl}, \pi(s^n)) \geq 1 - (n+1)^{|S|} 2^{-na} \tag{18}$$

for all $s^n \in S^n$. Hence by defining $\mathcal{C}^\pi := \{\pi^{-1}(x_{jl}^n), \pi^{-1}(D_j)\}$ as a member of a family of codes $\{\mathcal{C}^\pi\}_{\pi \in \Pi_n}$ together with a RV K distributed according to μ as the uniform distribution on Π_n, (18) is equivalent to

$$\mathbb{E}_\mu(\bar{\lambda}_n(\mathcal{C}^K, W_{s^n}^n)) \leq (n+1)^{|S|} 2^{-na'} =: \lambda_n \tag{19}$$

with $\bar{\lambda}_n(\mathcal{C}^\pi, W_{s^n}^n)$ as the respective average error probability for $K = \pi$ and it holds for all $s^n \in S^n$. Thus

$$\mathcal{C}_n^{\text{ran}} := \{(\pi^{-1}(x_{jl}), \pi^{-1}(D_j)) : j \in [J_n], l \in [L_n], \pi \in \Pi_n, \mu\}$$

is a random code for the AVC to the legitimate receiver with error probability bounded as in (19). Now it is easily seen

$$p_{JZ_{q^*}^n}^{\mathcal{C}^\pi}(j, z^n) = \frac{1}{J_n} \frac{1}{L_n} \sum_{l=1}^{L_n} V_{q^*}^{\otimes n}(\pi^{-1}(z^n) | \pi^{-1}(x_{jl})) = p_{JZ_{q^*}^n}(j, z^n) \ .$$

Then with the representation of the mutual information by the information divergence we obtain from (16)

$$\mathbb{E}_\mu(D(p_{JZ_{q^*}^n}^{\mathcal{C}^K} || p_J \otimes p_{Z_{q^*}^n}^{\mathcal{C}^K})) = \frac{1}{n!} \sum_{\pi \in \Pi_n} D(p_{JZ_{q^*}^n}^{\mathcal{C}^\pi} || p_J \otimes p_{Z_{q^*}^n}^{\mathcal{C}^\pi}) = I(p_J, V_{q^*}^{\otimes n}) \leq \epsilon' \ .$$

Thus we have constructed a random code $\mathcal{C}_n^{\text{ran}}$ with mean average error probability bounded for all $s^n \in S^n$ as in (19) and which fulfills the strong secrecy criterion, provided that there exist a best channel to the eavesdropper. By the construction of the random code it follows that the secrecy rates given by (15) for the compound wiretap channel $\overline{\mathcal{W}}$ achieved by the deterministic code $\mathcal{C}_{\overline{\mathcal{W}}}$ are achievable secrecy rates for the AVWC $\mathfrak{W}$ with random code $\mathcal{C}_n^{\text{ran}}$. ■

C. Deterministic Code Construction

Because the code $\mathcal{C}^\pi$ that is used for the transmission of a single message is subjected to a random selection, reliable transmission can only be guaranteed if the outcome of the random experiment is known to both the transmitter and the receiver. One way to inform the receiver about the code that is chosen is to add a short prefix to the actual codeword. Provided that the number of codes is small enough, the transmission of these additional prefixes causes no essential loss in rate. In the following we use the *elimination technique* of Ahlswede [7] who has introduced the above approach to derive deterministic codes from random codes for determining capacity of arbitrarily varying channels. Temporarily we drop the requirement of a best channel to the eavesdropper.

Theorem 3.7: 1) In the case that for the random code secrecy capacity of the AVWC $\mathfrak{W}$ it holds that $C_{S,\text{ran}}(\mathfrak{W}) > 0$ the secrecy capacity $C_S(\mathfrak{W})$ of the AVWC $\mathfrak{W}$ equals its random code capacity $C_{S,\text{ran}}(\mathfrak{W})$

$$C_S(\mathfrak{W}) = C_{S,\text{ran}}(\mathfrak{W}), \tag{20}$$

iff the channel to the legitimate receiver is non-symmetrisable. 2) If the channel to the legitimate receiver is non-symmetrisable it always holds that $C_S(\mathfrak{W}) = C_{S,\text{ran}}(\mathfrak{W})$.
If the channel to the legitimate receiver is symmetrisable then the deterministic code capacity of the channel to the legitimate receiver equals zero by Theorem 3.3. Hence the deterministic code secrecy capacity of the AVWC also equals zero although the random code secrecy capacity could be greater than zero. So we can restrict to the case in which the channel to the legitimate receiver is non-symmetrisable. If $C_S(\mathfrak{W}) = C_{S,\text{ran}}(\mathfrak{W}) > 0$ then the channel to the legitimate receiver must be nonsymmetrisable. For the other direction, because the secrecy capacity of the AVWC $\mathfrak{W}$ cannot be greater than the random code secrecy capacity it suffices to show that $C(\{W_{s^n}\}) > 0$ implies that $C_S(\mathfrak{W}) \geq C_{S,\text{ran}}(\mathfrak{W})$. Here $C(\{W_{s^n}\})$ denotes the capacity of the AVC to the legitimate receiver without secrecy. The proof is given in the two paragraphs *Random code reduction* and *Elimination of randomness*, cf. [11] for the AVC.

1) Random Code Reduction: We first reduce the random code $\mathcal{C}^{\text{ran}}$ to a new random code selecting only a small number of deterministic codes from the former, and averaging over this codes gives a random code with a constant small mean average error probability, which fulfills the secrecy criterion.

Lemma 3.8: (Random Code Reduction) Let $\mathcal{C}(\mathcal{Z})$ be a random code for the AVWC $\overline{\mathfrak{W}}$ consisting of a family $\{\mathcal{C}(\gamma)\}_{\gamma \in \Gamma}$ of wiretap codes where γ is chosen according to the distribution μ of $\mathcal{Z}$. Then let

$$\bar{e}(\mathcal{C}^{\text{ran}}) = \mathbb{E}_\mu \bar{e}(\mathcal{C}(\mathcal{Z})) \leq \lambda \text{ and } \max_{s^n} \mathbb{E}_\mu I(p_J, V_{s^n}; \mathcal{C}(\mathcal{Z})) \leq \epsilon'.$$

Then for any ϵ and K satisfying

$$\epsilon > 4 \max\{\lambda_n, \epsilon'\} \quad \text{and} \quad K > \frac{2n \log |A|}{\epsilon} (1 + n \log |S|) \tag{21}$$

there exist K deterministic codes $\mathcal{C}_i$, $i = 1, \dots, K$ chosen from the random code by random selection such that

$$\frac{1}{K}\sum_{i=1}^{K} \bar{e}(s^n|\mathcal{C}_i) \le \epsilon \quad \text{and} \quad \frac{1}{K}\sum_{i=1}^{K} I(p_J, V_{s^n}; \mathcal{C}_i) \le \epsilon \quad (22)$$

for all $s^n \in S^n$.

Proof: The proof is analogue to the proof of Lemma 6.8 [11], where the assertion for the maximal probability of error for ordinary AVC's is established. Let $\mathcal{Z}$ be the RV distributed according to μ on Γ for the random code. Now consider K independent repetitions of the random experiment of code selections and call the according RV $\mathcal{Z}_i$, $i \in \{1, \dots, K\}$. Then

$$\begin{aligned}
&\Pr\Big\{\frac{1}{K}\sum_{i=1}^{K}\bar{e}(s^n|\mathcal{C}(\mathcal{Z}_i)) \ge \epsilon \text{ or } \frac{1}{K}\sum_{i=1}^{K} I(p_J, V_{s^n}; \mathcal{C}(\mathcal{Z}_i)) \ge \epsilon\Big\} \\
&\le \exp\Big(-\frac{K\epsilon}{n\log|A|}\Big)\mathbb{E}\exp\sum_{i=1}^{K}\frac{\bar{e}(s^n|\mathcal{C}(\mathcal{Z}_i))}{n\log|A|} \\
&\qquad + \exp\Big(-\frac{K\epsilon}{n\log|A|}\Big)\mathbb{E}\exp\sum_{i=1}^{K}\frac{I(p_J, V_{s^n}; \mathcal{C}(\mathcal{Z}_i))}{n\log|A|},
\end{aligned}$$

holds for any $s^n \in S^n$, using Bernstein's trick and applying Markov's inequality. Now because of the independency of the RVs $\mathcal{Z}_i$ and they all are distributed as $\mathcal{Z}$ and we have $\exp t \le 1 + t$, for $0 \le t \le 1$ (exp to the base 2), we can give the following upper bounds

$$\Big(\mathbb{E}\exp\frac{\bar{e}(s^n|\mathcal{C}(\mathcal{Z}))}{n\log|A|}\Big)^K \le \Big(1 + \frac{\lambda_n}{n\log|A|}\Big)^K \ ,$$

$$\Big(\mathbb{E}\exp\frac{I(p_J, V_{s^n}; \mathcal{C}(\mathcal{Z}))}{n\log|A|}\Big)^K \le \Big(1 + \frac{\epsilon'}{n\log|A|}\Big)^K \ .$$

Hence, by applying the union bound we obtain

$$\begin{aligned}
&\Pr\Big\{\frac{1}{K}\sum_{i=1}^{K}\bar{e}(s^n|\mathcal{C}(\mathcal{Z}_i)) \le \epsilon \text{ and } \\
&\qquad \frac{1}{K}\sum_{i=1}^{K} I(p_J, V_{s^n}; \mathcal{C}(\mathcal{Z}_i)) \le \epsilon, \ \forall s^n \in S^n\Big\} \\
&\ge 1 - 2|S|^n \exp\Big[-K\Big(\frac{\epsilon}{n\log|A|} \\
&\qquad - \log\big(1 + \max\{\frac{\lambda_n}{n\log|A|}, \frac{\epsilon'}{n\log|A|}\}\big)\Big)\Big],
\end{aligned} \quad (23)$$

which is strictly positive, if we choose

$$\begin{aligned}
\epsilon &\ge 2n\log|A|\log(1 + \max\{\frac{\lambda_n}{n\log|A|}, \frac{\epsilon'}{n\log|A|}\}) \ , \\
K &\ge \frac{2\log|A|}{\epsilon}(n + n^2\log|S|) \ .
\end{aligned} \quad (24)$$

Now because for $0 \le t \le 1$ and log is to the base 2 it holds that $t \le \log(1 + t) \le 2t$, we increase the lower bound for choosing ϵ, if $\epsilon \ge 4\max\{\lambda_n, \epsilon'\}$ and with (24) the assertion of (23) still holds. Hence, we have shown that there exist K realisations $\mathcal{C}_i := \mathcal{C}(\mathcal{Z}_i = \gamma_i)$, $\gamma_i \in \Gamma$, $i \in \{1, \dots, K\}$ of the random code, which build a new reduced random code with uniform distribution on these codes with mean average error probability and mean secrecy criterion fulfilled by (22). ■

Now we assume that the channel to the legitimate receiver is non-symmetrisable, $C(\{W_{s^n}\}) > 0$, and there exist a random code $\mathcal{C}_n^{\text{ran}}$ that achieves the random code capacity $C_{S,\text{ran}}(\mathfrak{W}) > 0$. Then there exist a sequence of random codes with

$$\liminf_{n\to\infty}\frac{1}{n}\log J_n \to C_{S,\text{ran}}(\mathfrak{W}) > 0, \quad (25)$$

which fulfills the average error criterion and the strong secrecy criterion. Then on account of the random code reduction lemma there exist a sequence of random codes consisting only of n^3 deterministic codes (21) chosen from the former random code, and it holds for any $\epsilon > 0$ and sufficiently large n that

$$\max_{s^n\in S^n}\frac{1}{J_n}\sum_{j=1}^{J_n}\frac{1}{n^3}\sum_{i=1}^{n^3}\sum_{x^n\in A^n}E^i(x^n|j)W_{s_n}^{\otimes n}((D_j^i)^c|x^n) \le \epsilon$$

$$\text{and} \qquad \max_{s^n\in S^n}\frac{1}{n^3}\sum_{i=1}^{n^3} I(p_J; V_{s^n}^n; \mathcal{C}_i) \le \epsilon,$$

where $\mathcal{C}_i = \{(E_j^i, D_j^i), j \in \mathcal{J}_n\}$, $i = 1, \dots, n^3$, and E^i is the stochastic encoder of the deterministic wiretap code. Then the reduced random code consists of the family of codes $\{\mathcal{C}_i\}_{i\in\{1,\dots,n^3\}}$ together with the uniform distribution $\mu'(i) = \frac{1}{n^3}$ for all $i \in \{1, \dots, n^3\}$.

2) Elimination of randomness: (Cf. Theorem 6.11 in [11]) Now if there exist a deterministic code and $C(\{W_{s^n}\}) > 0$ then there exist a code $\{x_i^{k_n}, F_i \subset B^{k_n} : i = 1, \dots n^3\}$, $x_i^{k_n}$ is chosen according to an encoding function $f_i : \{1, \dots, n^3\} \to A^{k_n}$ with $\frac{k_n}{n} \to 0$ as $n \to \infty$, with error probability

$$\frac{1}{n^3}\sum_{i=1}^{n^3} W^{\otimes k_n}(F_i^c|x_i^{k_n}, s^{k_n}) \le \epsilon \qquad \text{for all } s^{k_n} \in S^{k_n} \ .$$

If we now compose a new deterministic code for the AVWC $\mathfrak{W}$ by prefixing the codewords of each C_i

$$\{f_i E_j^i, F_i \times D_j^i : i = 1, \dots, n^3, j \in [J_n]\} =: \mathcal{C} \ ,$$

the decoder is informed of which encoder E^i is in use for the actual message j if he identifies the prefix correctly. Note that for the transmission of the prefix only the reliability is of interest. Now the new codewords have a length of $k_n + n$, transmitting a message from $\{1, \dots, n^3\} \times \mathcal{J}_n$ via the channel with state sequence s^{k_n+n} yields an average error probability

$$\bar{\lambda}_n(\mathcal{C}, W_{s^{k_n+n}}^{\otimes(k_n+n)}) \le \frac{1}{n^3 J_n}\sum_{i=1}^{n^3}\sum_{j\in[J_n]}(\lambda_i + \lambda_j(i)) \le 2\epsilon. \quad (26)$$

Here, for each $s^{k_n} \in S^{k_n}$ λ_i denotes the error probability for transmitting i encoded in $x_i^{k_n}$ by $W_{s^{k_n}}^{k_n}$ followed by the transmission of j, where the codeword is chosen according to the stochastic encoder E_j^i, over the last n channel realisations determined by s^n with error probability $\lambda_j(i)$. This construction is possible due to the memorylessness of the channel.

Now if we turn to the security part it is easily seen that

$$p_{JZ_{s^{k_n+n}}^{k_n+n}}^{\mathcal{C}}(j, z^{k_n+n}) = \frac{1}{n^3}\sum_{i=1}^{n^3} V_{s^{k_n}}^{\otimes k_n}(\hat{z}^{k_n}|x_i^{k_n}) \cdot p_{JZ_{s^n}^n}^{\mathcal{C}_i}(j, z^n) \ ,$$

which is caused by the memorylessness of the channel. Here $\hat{z}^{k_n}$ denotes the first k_n components of z^{k_n+n}. By using this equation and the representation of the mutual information by the informational divergence we obtain that

$$\begin{aligned} &D(p^{\mathcal{C}}_{JZ^{k_n+n}_{s^{k_n+n}}} || p_J \otimes p^{\mathcal{C}}_{Z^{k_n+n}_{s^{k_n+n}}}) \\ &\leq \frac{1}{n^3}\sum_{i=1}^{n^3} D\big(V^{k_n}_{s^{k_n}}(\hat{z}^{k_n}|x_i^{k_n}) p^{\mathcal{C}_i}_{JZ^n_{s^n}} \big\| V^{k_n}_{s^{k_n}}(\hat{z}^{k_n}|x_i^{k_n}) p_J \otimes p^{\mathcal{C}_i}_{Z^n_{s^n}}\big) \\ &= \frac{1}{n^3}\sum_{i=1}^{n^3} D\big(p^{\mathcal{C}_i}_{JZ^n_{s^n}} \big\| p_J \otimes p^{\mathcal{C}_i}_{Z^n_{s^n}}\big) = \frac{1}{n^3}\sum_{i=1}^{n^3} I(p_J, V^n_{s^n}; \mathcal{C}_i) \leq \epsilon \end{aligned} \tag{27}$$

for all $s^n \in S^n$ and $n \in \mathbb{N}$ sufficiently large, where the first inequality follows because the relative entropy $D(p\|q)$ is a convex function in the pair (p,q) and the last inequality follows by the random code reduction lemma.

Because $\frac{k_n}{n} \to 0$ as $n \to \infty$

$$\lim_{n\to\infty} \frac{1}{k_n+n}\log(n^3 J_n) = \lim_{n\to\infty}\frac{1}{n}\log J_n \ . \tag{28}$$

Thus, with $\{1,\ldots,J_n\}$ as the message set, $\mathcal{C}_n$ is a deterministic $(n+o(n), n^3\cdot J_n)$ code with average error probability bounded for all $s^{k_n+n} \in S^{k_n+n}$ as in (26) and which fulfills the strong secrecy criterion as in (27). By (28) it achieves the random code secrecy capacity $C_{S,\text{ran}}$ of the AVWC $\mathfrak{W}$ (25), which implies that $C_S = C_{S,\text{ran}}$.

As a consequence of Theorem 3.7 and Proposition 3.6 we can state the following assertion.

Corollary 3.9: Provided that there exists a best channel to the eavesdropper and under the assumption that the channel to the legitimate receiver is non-symmetrisable, the deterministic code secrecy capacity of the AVWC $\mathfrak{W}$ is lower bounded by

$$C_S(\mathfrak{W}) \geq \max_{p\in\mathcal{P}(A)}\big(\min_{q\in\mathcal{P}(S)} I(p,W_q) - \max_{q\in\mathcal{P}(S)} I(p,V_q)\big) \ .$$

D. An upper bound on the capacity of the AVWC $\mathfrak{W}$ and a multiletter coding theorem

First we give an upper bound on the secrecy capacity of the AVWC $\mathfrak{W}$ which corresponds to the bound for the compound wiretap channel built by the same family of channels.

Theorem 3.10: The secrecy capacity of the AVWC $\mathfrak{W}$ is upper bounded,

$$C_S(\mathfrak{W}) \leq \min_{q\in\mathcal{P}(S)} \max_{U\to X\to(YZ)_q} (I(U,Y_q) - I(U,Z_q)) \ . \tag{29}$$

Proof: By Lemma 3.4 the capacity of the AVWC $\mathfrak{W}$ equals the capacity of the AVWC $\overline{\mathfrak{W}}$. Obviously, the set $\overline{\mathcal{W}} = \{(W_q^{\otimes n}, V_q^{\otimes n}) : q \in \mathcal{P}(S)\}$ which describes a compound wiretap channel is a subset of $\overline{\mathfrak{W}}^n = \{(W^n_{\tilde{q}}, V^n_{\tilde{q}}) : \tilde{q} \in \mathcal{P}(S^n), \tilde{q} = \prod_{i=1}^n q_i\}$. Then it follows that

$$C_S(\mathfrak{W}) \leq \inf_{\tilde{q}} C_S((W^n_{\tilde{q}}, V^n_{\tilde{q}})) \leq \inf_{q} C_S((W^n_q, V^n_q)).$$

The minimum is attained because of the continuity of $C_S(W_q, V_q)$ on the compact set $\overline{\mathfrak{W}}$. ∎

Remark 3.11: Consider the special case of an AVWC $\mathfrak{W} = \{(W_{s^n}, V_{r^n}) : s^n \in S_1^n,\ r^n \in S_2^n\}$, where the states of both the channels in every time step can be chosen independently. In addition let us assume that there exist a channel $W_{q_1^*} \in \{W_{q_1} : q_1 \in \mathcal{P}(S_1)\}$, which is a degraded version of the other channels in the set, and a best channel to the eavesdropper $V_{q_2^*}$ from the set $\{V_{q_2} : q_2 \in \mathcal{P}(S_2)\}$. Then the lower bound on the secrecy capacity given in Corollary 3.9 matches the upper bound from Theorem 3.10 [2]. Thus we can conclude that under the assumption, that the channel to legitimate receiver is non-symmetrisable, the capacity of the AVWC $\mathfrak{W}$ is given by

$$C_S(\mathfrak{W}) = \max_{p\in\mathcal{P}(A)} (I(p,W_{q_1^*}) - I(p,V_{q_2^*})) \ .$$

Now we give a multiletter upper bound on the secrecy rates. In the sense of Lemma 3.4 "good" codes for the AVWC $\mathfrak{W}$ are "good" codes for the AVWC $\overline{\mathfrak{W}}$. Now because $\overline{\mathcal{W}} = \{(W_q^{\otimes n}, V_q^{\otimes n}) : q \in \mathcal{P}(S)\}$ is a subset of $\overline{\mathfrak{W}}^n$ those codes are "good" codes for the compound channel $\overline{\mathcal{W}}$. Thus the secrecy capacity $C_S(\mathfrak{W})$ of the AVWC $\mathfrak{W}$ is upper bounded by

$$\lim_{n\to\infty}\frac{1}{n}\max_{U\to X^n\to(Y^nZ^n)_q}\big(\inf_{q\in\mathcal{P}(S)} I(U,Y_q^n) - \sup_{q\in\mathcal{P}(S)} I(U,Z_q^n)\big)$$

by Proposition 3.8 in [2]. Under the assumption that the channel to the legitimate receiver is non-symmetrisable and provided that there exist a best channel to the eavesdropper this term equals the secrecy capacity $C_S(\mathfrak{W})$, which is easily seen if we apply the assertion of Corollary 3.9 to the n-fold product of channels W_q and V_q.

ACKNOWLEDGMENT

Support by the Deutsche Forschungsgemeinschaft (DFG) via projects BO 1734/16-1, BO 1734/20-1, and by the Bundesministerium für Bildung und Forschung (BMBF) via grant 01BQ1050 is gratefully acknowledged.

REFERENCES

[1] I. Bjelaković, H. Boche, and J. Sommerfeld, "Capacity results for compound wiretap channels," *Proc. IEEE Information Theory Workshop*, pp. 60–64, 2011.

[2] ——, "Secrecy results for compound wiretap channels," 2011, submitted to Problems of Information Transmission. [Online]. Available: http://arxiv.org/abs/1106.2013v1

[3] Y. Liang, G. Kramer, H. Poor, and S. Shamai, "Compound Wiretap Channels," *EURASIP Journal on Wireless Communications and Networking*, 2008.

[4] M. Bloch and J. Laneman, "On the secrecy capacity of arbitrary wiretap channel," *Forty-Sixth Annual Allerton Conference, Allerton House, Illinois, USA*, Sep. 2008.

[5] R. Ahlswede, "Arbitrarily varying channels with states sequence known to the sender," *IEEE Transactions on Information Theory*, vol. 32, no. 5, pp. 621–629, Sept 1986.

[6] E. MolavianJazi, M. Bloch, and J. Laneman, "Arbitrary jamming can preclude secure communications," *Proc. 47th Annual Allerton Conference on Communication, Control, and Computing, Monticello, IL*, pp. 1069–1075, 2009.

[7] R. Ahlswede, "Elimination of correlation in random codes for arbitrarily varying channels," *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, vol. 44, pp. 159–175, 1978.

[8] T. Ericson, "Exponential error bounds for random codes in the arbitrarily varying channel," *IEEE Transactions on Information Theory*, vol. 31, no. 1, pp. 42–48, 1985.

[9] I. Csiszar and P. Narayan, "The capacity of the arbitrarily varying channel revisited: Positivity, constraints," *IEEE Transactions on Information Theory*, vol. 34, no. 2, pp. 181–193, 1988.

[10] E. MolavianJazi, "Secure communications over arbitrarily varying wiretap channels," Master's thesis, Graduate School of the University of Notre Dame, 2009.

[11] I. Csiszar and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Akademiai Kiado, 1981.