Characterization of Super-Additivity and Discontinuity Behavior of the Capacity of Arbitrarily Varying Channels under List Decoding

Holger Boche
Institute of Theoretical Information Technology
Technische Universität München

Rafael F. Schaefer
Information Theory and Applications Chair
Technische Universität Berlin

H. Vincent Poor
Dept. of Electrical Engineering
Princeton University

***Abstract*—The *arbitrarily varying channel (AVC)* models communication over a channel that varies in an arbitrary and unknown manner from channel use to channel use. This paper considers the AVC under list decoding and studies the corresponding list capacity. In particular, the list capacity function is shown to be discontinuous and the corresponding discontinuity points are characterized for all possible list sizes. For orthogonal AVCs it is then shown that the list capacity is super-additive, implying that joint encoding and decoding for two orthogonal AVCs can yield a larger list capacity than independent processing of both channels. This discrepancy is shown to be arbitrary large.**

I. Introduction

Arbitrarily varying channels (AVCs) [1–3] model communication with imperfect channel state information. This concept assumes that the actual channel realization is unknown; it is only known that this realization is from a known uncertainty set and that it varies in an arbitrary and unknown manner from channel use to channel use. This framework not only models the case of channel uncertainty, but also captures scenarios with interference from malevolent adversaries.

Reliable communication over AVCs is a non-trivial task; in particular, for so-called *symmetrizable* AVCs it has been shown that traditional coding schemes with pre-specified encoder and decoder are not sufficient [2, 3]. More sophisticated coding schemes based on common randomness (CR) are needed [1]. Such coding schemes might not be feasible in practice when the transmitter and receiver have no access to such coordination resources so that list decoding has been proposed to overcome such problems. The AVC under list decoding has been studied in [4] and [5]. In particular, it has been shown that the list capacity of an AVC displays a dichotomous behavior: Whenever the list size is larger than the symmetrizability of the channel, the list capacity equals the CR-assisted capacity; otherwise it is zero. In this paper, we further study the behavior of the list capacity function and show that it is discontinuous, and present a complete characterization of this behavior. This means that small variations in the uncertainty set can lead to dramatic changes in the list capacity.

We continue our study by addressing the question of additivity of the list capacity function. Intuitively, a capacity function is expected to be additive: Given two orthogonal channels the capacity of the overall system should be the sum of the individual channels. This is indeed the case for discrete memoryless channels but in general by no means trivial or obvious to answer. Shannon for example asked this question in 1956 for the zero error capacity [6] and conjectured it to be additive as well. This was disproved by Haemers [7] and Alon [8] by constructing counter-examples. To date, a general characterization of this phenomenon is not known and only certain explicit examples of super-additivity are known. Alon further conjectured that the additivity is even violated in a strong form. This was recently disproved by Keevash who showed in [9] that the discrepancy of the normalized Shannon zero error capacity between joint and independent processing of orthogonal channels is indeed bounded.

In another line of research, Ahlswede showed that the characterization of the zero error capacity is included as a special case in the problem of determining the capacity of the AVC under the maximal error criterion [10]. This connects these two fields making it worth studying the question of additivity also from an AVC perspective. This has been done for the AVC in [11] where it has been shown that the deterministic capacity of an AVC under the average error criterion is super-additive including a complete characterization. In this paper, we extend these studies to the AVC under list decoding and show that the list capacity is super-additive as well. In addition, we study whether the strong form of violation according to Alon's conjecture happens for the list capacity as well, and show that the discrepancy in list capacity between a joint use of two orthogonal AVCs and its corresponding independent use can be arbitrarily large. To achieve this, a unified theory is developed that allows us to answer such questions of discontinuity and non-additivity not only for the list capacity as done in this paper, but also for the capacity with maximum error and randomized encoding, as well as for the ϵ-capacity with average error. It is an interesting and open question to further extend these studies to the AVC with state constraints.[1]

This research was supported in part by the German Research Foundation (DFG) under Grant BO 1734/20-1 and in part by the U. S. National Science Foundation under Grants CMMI-1435778 and ECCS-1647198.

[1] *Notation:* $\mathcal{P}(\cdot)$ denotes the set of all probability distributions on its argument; $\mathrm{Sym}[L]$ denotes the set of all permutations on $\{1, 2, ..., L\}$; $\mathrm{CH}(\mathcal{X}; \mathcal{Y})$ denotes the set of all stochastic matrices (channels) $\mathcal{X} \to \mathcal{P}(\mathcal{Y})$.

978-1-5090-4096-4/17/$31.00 ©2017 IEEE

II. Arbitrarily Varying Channel

Let $\mathcal{X}$ and $\mathcal{Y}$ be finite input and output sets and $\mathcal{S}$ be a finite state set. Then the channel between the transmitter and the receiver is given by a stochastic matrix $W : \mathcal{X} \times \mathcal{S} \to \mathcal{P}(\mathcal{Y})$ which we interchangeably also write as $W_{\mathcal{S}} : \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ and $W \in \mathrm{CH}(\mathcal{X}, \mathcal{S}; \mathcal{Y})$. For a fixed state sequence $s^n \in \mathcal{S}^n$, the discrete memoryless channel is given by $W^n_{s^n}(y^n|x^n) = W^n(y^n|x^n, s^n) = \prod_{i=1}^n W(y_i|x_i, s_i)$ for all input and output sequences $x^n \in \mathcal{X}^n$ and $y^n \in \mathcal{Y}^n$.

Definition 1. The *arbitrarily varying channel (AVC)* $\mathfrak{W}$ is given by

$$\mathfrak{W} = \{W_s\}_{s\in\mathcal{S}} = \{W(\cdot|\cdot, s)\}_{s\in\mathcal{S}}.$$

If not otherwise stated, all AVCs are assumed to be *finite*. This means that the input and output alphabets $\mathcal{X}$ and $\mathcal{Y}$ are finite and, in particular, that the state set $\mathcal{S}$ is finite.

A. List Codes

For AVCs it makes a substantial difference what kind of codes are used for communication and whether additional resources for coordination are available or not [1–5].

1) List Codes: Instead of decoding the received signal into exactly one message, the decoder of a list code outputs a list of up to L possible messages.

Definition 2. A (n, M_n)-*list-L code* $\mathcal{C}_L$ consists of a deterministic encoder at the transmitter $f : \mathcal{M} \to \mathcal{X}^n$ with a set of messages $\mathcal{M} = \{1, ..., M_n\}$ and a list decoder at the receiver $\varphi_L : \mathcal{Y}^n \to \mathfrak{P}_L(\mathcal{M})$ with $\mathfrak{P}_L(\mathcal{M})$ the set of all subsets of $\mathcal{M}$ with cardinality at most L.

For a state sequence $s^n \in \mathcal{S}^n$ the average probability of error of such a list code is

$$\bar{e}_L(s^n) = \frac{1}{|\mathcal{M}|} \sum_{m\in\mathcal{M}} \sum_{y^n : m\notin\varphi_L(y^n)} W^n(y^n|x^n_m, s^n).$$

Definition 3. A rate $R > 0$ is an *achievable list-L rate* for an AVC $\mathfrak{W}$ if for all $\tau > 0$ and there exists an $n(\tau) \in \mathbb{N}$ and a sequence $\{\mathcal{C}_{L,n}\}_{n\in\mathbb{N}}$ such that for all $n \geq n(\tau)$ we have

$$\frac{1}{n} \log \frac{M_n}{L} \geq R - \tau \quad \text{and} \quad \max_{s^n\in\mathcal{S}^n} \bar{e}_L(s^n) \leq \lambda_n$$

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

The *list-L capacity* $C_L(\mathfrak{W})$ of an AVC $\mathfrak{W}$ is given by the supremum of all achievable list-L rates R.

Remark 1. For list size $L = 1$ the list-L code $\mathcal{C}_L$ in Definition 2 reduces to a traditional deterministic code $\mathcal{C}$ whose decoder outputs only one specific message, i.e., $\varphi : \mathcal{Y}^n \to \mathcal{M}$.

2) CR-Assisted Codes: CR is modeled by a random variable Γ taking values in a finite set $\mathcal{G}_n$ according to a probability distribution $P_\Gamma \in \mathcal{P}(\mathcal{G}_n)$. This enables the transmitter and receiver to coordinate their choice of encoder and decoder based on the actual realization $\gamma \in \mathcal{G}_n$.

Definition 4. A *CR-assisted* $(n, M_n, \mathcal{G}_n, P_\Gamma)$*-code* $\mathcal{C}_{\mathrm{CR}}$ is given by a family of deterministic codes $\{\mathcal{C}(\gamma) : \gamma \in \mathcal{G}_n\}$ together with a random variable Γ taking values in $\mathcal{G}_n$ according to $P_\Gamma \in \mathcal{P}(\mathcal{G}_n)$.

The average error extends to CR-assisted codes by taking the expectation over the family of codes. The definitions of a *CR-assisted achievable rate* and the *CR-assisted capacity* $C_{\mathrm{CR}}(\mathfrak{W})$ of an AVC $\mathfrak{W}$ follow accordingly.

B. Capacity Results

The CR-assisted capacity of an AVC goes back to [1] and is restated next.

Theorem 1 ([1]). The CR-assisted capacity $C_{CR}(\mathfrak{W})$ of an AVC $\mathfrak{W}$ is

$$C_{CR}(\mathfrak{W}) = \max_{P_X\in\mathcal{P}(\mathcal{X})} \inf_{q\in\mathcal{P}(\mathcal{S})} I(X; \overline{Y}_q)$$

where $\overline{Y}_q$ represents the output of the averaged channel $\overline{W}_q(y|x) = \sum_{s\in\mathcal{S}} W(y|x, s)q(s)$ for some $q \in \mathcal{P}(\mathcal{S})$.

To characterize the list-L capacity of an AVC, we need the concept of symmetrizability which, roughly speaking, describes the ability of an AVC to "simulate" additional valid inputs making it impossible for the decoder to decide on the correct codeword.

Definition 5. An AVC is called *L-symmetrizable* if there exists a stochastic matrix $\sigma \in \mathrm{CH}(\mathcal{X}^L; \mathcal{S})$ such that for all permutations $\pi \in \mathrm{Sym}[L + 1]$ the following holds:

$$\sum_{s\in\mathcal{S}} W(y|x_1, s)\sigma(s|x_2, ..., x_{L+1}) = \sum_{s\in\mathcal{S}} W(y|x_{\pi(1)}, s)\sigma(s|x_{\pi(2)}, ..., x_{\pi(L+1)})$$

for all $x_1, x_2, ..., x_{L+1} \in \mathcal{X}$ and $y \in \mathcal{Y}$.

With this the list-L capacity is characterized as follows.

Theorem 2 ([4, 5]). The list-L capacity $C_L(\mathfrak{W})$ of an AVC $\mathfrak{W}$ is

$$C_L(\mathfrak{W}) = \begin{cases} C_{CR}(\mathfrak{W}) & \text{if } \mathfrak{W} \text{ is not } L\text{-symmetrizable} \\ 0 & \text{otherwise}. \end{cases}$$

III. Discontinuity and Super-Additivity Behavior

Here, we further explore the concept of list decoding for AVCs and identify certain properties of the list-L capacity.

A. A Fundamental Function and Basic Properties

We introduce the function $F_L(\mathfrak{W}) : \mathrm{CH}(\mathcal{X}, \mathcal{S}; \mathcal{Y}) \to \mathbb{R}_+$

$$F_L(\mathfrak{W}) = \inf_{\sigma\in\mathrm{CH}(\mathcal{X}^L;\mathcal{S})} \max_{x^{L+1}\in\mathcal{X}^{L+1}} \max_{\pi\in\mathrm{Sym}[L+1]} \sum_{y\in\mathcal{Y}} \Big| \sum_{s\in\mathcal{S}} W(y|x_1, s)\sigma(s|x_2, ..., x_{L+1}) - \sum_{s\in\mathcal{S}} W(y|x_{\pi(1)}, s)\sigma(s|x_{\pi(2)}, ..., x_{\pi(L+1)}) \Big|.$$

Since $\mathrm{CH}(\mathcal{X}^L; \mathcal{S})$ is a bounded and closed set, there exists for any AVC $\mathfrak{W}$ a channel $\sigma^* \in \mathrm{CH}(\mathcal{X}^L; \mathcal{S})$ such that the infimum above is achieved and F_L can be expressed as a minimum.

Further, we have $F_L(\mathfrak{W}) \geq 0$ with equality if and only if $\mathfrak{W}$ is L-symmetrizable; cf. Definition 5.

We also need a concept of distance. For two DMCs $W_1, W_2 \in \mathrm{CH}(\mathcal{X};\mathcal{Y})$ we define the distance between W_1 and W_2 based on the total variation distance as

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

To extend this concept to AVCs, we consider two AVCs $\mathfrak{W}_1 = \{W_1(\cdot|\cdot,s_1)\}_{s_1\in\mathcal{S}_1}$ and $\mathfrak{W}_2 = \{W_2(\cdot|\cdot,s_2)\}_{s_2\in\mathcal{S}_2}$ with $W_i(\cdot|\cdot,s_i) \in \mathrm{CH}(\mathcal{X},\mathcal{S}_i;\mathcal{Y})$, $i = 1,2$, and define

$$G(\mathfrak{W}_1,\mathfrak{W}_2) := \max_{s_1\in\mathcal{S}_1} \min_{s_2\in\mathcal{S}_2} d\big(W_1(\cdot|\cdot,s_1), W_2(\cdot|\cdot,s_2)\big)$$

which describes how well one AVC can be approximated by the other one. Note that the function G is not symmetric. Accordingly, we define the distance between $\mathfrak{W}_1$ and $\mathfrak{W}_2$ as

$$D(\mathfrak{W}_1,\mathfrak{W}_2) := \max\big\{G(\mathfrak{W}_1,\mathfrak{W}_2), G(\mathfrak{W}_2,\mathfrak{W}_1)\big\}. \quad (1)$$

Note that $\mathcal{S}_1$ and $\mathcal{S}_2$ can be arbitrary finite state sets and we do not need $|\mathcal{S}_1| = |\mathcal{S}_2|$.

Lemma 1. The following inequalities hold:

$$F_L(\mathfrak{W}_2) \leq 2G(\mathfrak{W}_1,\mathfrak{W}_2) + F_L(\mathfrak{W}_1) \quad (2a)$$
$$F_L(\mathfrak{W}_1) \leq 2G(\mathfrak{W}_2,\mathfrak{W}_1) + F_L(\mathfrak{W}_2) \quad (2b)$$
$$\big|F_L(\mathfrak{W}_1) - F_L(\mathfrak{W}_2)\big| \leq 2D(\mathfrak{W}_1,\mathfrak{W}_2). \quad (2c)$$

Proof: The proof is omitted due to space constraints. ■

As an immediate consequence, we obtain the following.

Lemma 2. Let $\hat{\mathfrak{W}}$ be an arbitrary finite AVC and let $\{\mathfrak{W}_n\}_{n=1}^\infty$ be a sequence of finite AVCs such that

$$\lim_{n\to\infty} D(\mathfrak{W}_n,\hat{\mathfrak{W}}) = 0$$

holds. Then

$$\lim_{n\to\infty} F_L(\mathfrak{W}_n) = F_L(\hat{\mathfrak{W}}).$$

Proof: This follows immediately from Lemma 1. ■

For the next result, we need the concept of orthogonal (or parallel) AVCs. For two AVCs $\mathfrak{W}_1$ and $\mathfrak{W}_2$ we define the orthogonal AVC $\mathfrak{W}$ as

$$\mathfrak{W} = \mathfrak{W}_1 \otimes \mathfrak{W}_2 = \{W_1(\cdot|\cdot,s_1)\}_{s_1\in\mathcal{S}_1} \otimes \{W_2(\cdot|\cdot,s_2)\}_{s_2\in\mathcal{S}_2} \quad (3)$$

which means that the underlying channel law is

$$W(y_1,y_2|x_1,x_2,s_1,s_2) = W_1(y_1|x_1,s_1)W_2(y_2|x_2,s_2) \quad (4)$$

for all $x_i \in \mathcal{X}_i$, $y_i \in \mathcal{Y}_i$, and $s_i \in \mathcal{S}_i$, $i = 1,2$. Then we obtain the following upper and lower bounds on the function F_L.

Lemma 3. The following chain of inequalities holds:

$$\max\big\{F_L(\mathfrak{W}_1), F_L(\mathfrak{W}_2)\big\} \leq F_L(\mathfrak{W}_1 \otimes \mathfrak{W}_2) \leq F_L(\mathfrak{W}_1) + F_L(\mathfrak{W}_2).$$

Proof: The proof is omitted due to space constraints. ■

B. Characterization of Discontinuity Behavior

Here, we want to study the continuity of the list-L capacity, which we define as follows.

Definition 6. Let $\mathfrak{W}$ be a finite AVC. The list-L capacity $C_L(\mathfrak{W})$ is said to be *continuous* in $\mathfrak{W}$, if for all sequences of finite AVCs $\{\mathfrak{W}_n\}_{n=1}^\infty$ with

$$\lim_{n\to\infty} D(\mathfrak{W}_n,\mathfrak{W}) = 0 \quad (5)$$

we have

$$\lim_{n\to\infty} C_L(\mathfrak{W}_n) = C_L(\mathfrak{W}).$$

Based on this definition, the list-L capacity $C_L(\mathfrak{W})$ is discontinuous in $\mathfrak{W}$ if and only if there is a sequence $\{\mathfrak{W}_n\}_{n=1}^\infty$ of finite AVCs satisfying (5) but

$$\limsup_{n\to\infty} C_L(\mathfrak{W}_n) > \liminf_{n\to\infty} C_L(\mathfrak{W}_n)$$

is satisfied. With this concept, we are now in the position to give a complete characterization of the discontinuity points of the list-L capacity.

Theorem 3. The list-L capacity $C_L(\mathfrak{W})$ is discontinuous in $\mathfrak{W}$ if and only if the following conditions hold:

1) *$C_{CR}(\mathfrak{W}) > 0$*
2) *$F_L(\mathfrak{W}) = 0$ and for every $\epsilon > 0$ there exists a finite AVC $\hat{\mathfrak{W}}$ with $D(\mathfrak{W},\hat{\mathfrak{W}}) < \epsilon$ and $F_L(\hat{\mathfrak{W}}) > 0$.*

Proof: "⇒" First we show that both conditions are necessary for the list capacity to be discontinuous. We start with the first condition and assume that $\mathfrak{W}$ is a discontinuity point of $C_L(\mathfrak{W})$. Then there must exist a sequence of finite AVCs $\{\mathfrak{W}_n\}_{n=1}^\infty$ such that

$$\lim_{n\to\infty} D(\mathfrak{W}_n,\mathfrak{W}) = 0 \quad (6)$$

and

$$\limsup_{n\to\infty} \big|C_L(\mathfrak{W}_n) - C_L(\mathfrak{W})\big| > 0.$$

Then the first condition $C_{\mathrm{CR}}(\mathfrak{W}) > 0$ must hold due to the following reasoning: If $C_{\mathrm{CR}}(\mathfrak{W}) = 0$, we would have $C_L(\mathfrak{W}) = C_{\mathrm{CR}}(\mathfrak{W}) = 0$ as well. Further, for any sequence of finite AVCs $\{\mathfrak{W}_n\}_{n=1}^\infty$ for which (6) is satisfied we have $0 \leq C_L(\mathfrak{W}_n) \leq C_{\mathrm{CR}}(\mathfrak{W}_n)$. It is easy to see that the CR-assisted capacity $C_{\mathrm{CR}}(\mathfrak{W})$ is continuous in $\mathfrak{W}$, so that we would have $\lim_{n\to\infty} C_{\mathrm{CR}}(\mathfrak{W}_n) = 0$ and therewith also $\lim_{n\to\infty} C_L(\mathfrak{W}_n) = 0$, i.e., the list capacity would be continuous in $\mathfrak{W}$. But this contradicts our initial assumption so that the first condition $C_{\mathrm{CR}}(\mathfrak{W}) > 0$ must hold.

Next, we show that the second condition must be satisfied as well. If we would have $F_L(\mathfrak{W}) > 0$, the continuity of $F_L(\cdot)$ would imply $F_L(\hat{\mathfrak{W}}) > 0$ for all finite AVCs $\hat{\mathfrak{W}}$ with $D(\hat{\mathfrak{W}},\mathfrak{W}) < \epsilon$ for a suitable $\epsilon > 0$. Then for all finite AVCs $\hat{\mathfrak{W}}$ we would have $C_L(\hat{\mathfrak{W}}) = C_{\mathrm{CR}}(\hat{\mathfrak{W}})$ so that $C_L(\mathfrak{W})$ is continuous in $\mathfrak{W}$ which is a contradiction. Accordingly, we must have $F_L(\mathfrak{W}) = 0$. In the following we assume that the second condition is not satisfied, i.e., there exists an $\hat{\epsilon} > 0$ such that for all finite AVCs $\hat{\mathfrak{W}}$ with $D(\mathfrak{W},\hat{\mathfrak{W}}) < \hat{\epsilon}$ we have

$F_L(\hat{\mathfrak{W}})=0$. This implies that $\hat{\mathfrak{W}}$ is L-symmetrizable so that $C_L(\hat{\mathfrak{W}})=0$, cf. Theorem 2. This is true for all finite AVCs $\hat{\mathfrak{W}}$ that are close to $\mathfrak{W}$ (in the sense that $D(\mathfrak{W},\hat{\mathfrak{W}})<\hat{\epsilon}$) which means that $C_L(\mathfrak{W})$ is actually constant in the $\hat{\epsilon}$-neighborhood of $\mathfrak{W}$. Thus, it is continuous and the second condition must be satisfied as well.

"$\Leftarrow$" Now we prove the other direction and show that both conditions are also sufficient for the list-L capacity to be discontinuous. Since $F_L(\mathfrak{W})=0$ we have $C_L(\mathfrak{W})=0$; see also the initial discussion about F_L above. We choose a sequence of finite AVCs $\{\mathfrak{W}_n\}_{n=1}^{\infty}$ that satisfy $D(\mathfrak{W}_n,\mathfrak{W})<\frac{1}{n}$ and $F_L(\mathfrak{W}_n)>0$. With this choice, we have for all $n\in\mathbb{N}$ $C_L(\mathfrak{W}_n)=C_{\mathrm{CR}}(\mathfrak{W}_n)$ and further

$$\lim_{n\to\infty}C_L(\mathfrak{W}_n)=\lim_{n\to\infty}C_{\mathrm{CR}}(\mathfrak{W}_n)=C_{\mathrm{CR}}(\mathfrak{W})>0\neq C_L(\mathfrak{W})$$

so that $C_L(\mathfrak{W})$ is discontinuous in $\mathfrak{W}$. ∎

The next result establishes certain robustness properties of the list-L capacity.

Theorem 4. Let $\mathfrak{W}$ be a finite AVC with $F_L(\mathfrak{W})>0$. Then there exists an $\hat{\epsilon}>0$ such that all finite AVCs $\hat{\mathfrak{W}}$ with $D(\hat{\mathfrak{W}},\mathfrak{W})<\hat{\epsilon}$ are continuity points of $C_L(\mathfrak{W})$.

Proof: We have $F_L(\mathfrak{W})>0$. Since $F_L(\mathfrak{W})$ is a continuous function in $\mathfrak{W}$, cf. also Lemma 2, there exists an $\hat{\epsilon}$ such that for all finite AVCs $\hat{\mathfrak{W}}$ with $D(\hat{\mathfrak{W}},\mathfrak{W})<\hat{\epsilon}$ we always have $F_L(\hat{\mathfrak{W}})>0$. This implies $C_L(\hat{\mathfrak{W}})=C_{\mathrm{CR}}(\hat{\mathfrak{W}})$. Since $C_{\mathrm{CR}}(\mathfrak{W})$ is continuous in $\mathfrak{W}$, all these AVCs are continuity points of $C_L(\mathfrak{W})$. ∎

C. Super-Additivity for Orthogonal AVCs

Here, we study communication over two orthogonal AVCs and identify properties of the list-L capacity.

Definition 7. Let $\mathfrak{W}_1$ and $\mathfrak{W}_2$ be two finite AVCs and $\mathfrak{W}_1\otimes\mathfrak{W}_2$ an orthogonal combination as defined in (3) and (4). Then, the list-L capacity is said to be *super-additive* if

$$C_L(\mathfrak{W}_1\otimes\mathfrak{W}_2)>C_L(\mathfrak{W}_1)+C_L(\mathfrak{W}_2),$$

i.e., a joint use of both channels yields a higher list-L capacity than the sum of their individual uses.

Furthermore, if $C_L(\mathfrak{W}_1)=C_L(\mathfrak{W}_2)=0$ but $C_L(\mathfrak{W}_1\otimes\mathfrak{W}_2)>0$, we have the extreme case of non-additivity which we call *super-activation* of $\mathfrak{W}_1\otimes\mathfrak{W}_2$.

With this we can study the list-L capacity and show that super-activation is not possible for orthogonal AVCs.

Theorem 5. Let $\mathfrak{W}_1$ and $\mathfrak{W}_2$ be two orthogonal AVCs. Then

$$C_L(\mathfrak{W}_1\otimes\mathfrak{W}_2)=0 \tag{7}$$

if and only if

$$C_L(\mathfrak{W}_1)=C_L(\mathfrak{W}_2)=0. \tag{8}$$

Proof: "$\Leftarrow$" First, we show that (8) implies (7) with the help of Lemma 3. We have (8) if and only if

$$F_L(\mathfrak{W}_1)=F_L(\mathfrak{W}_2)=0. \tag{9}$$

Note that $C_{\mathrm{CR}}(\mathfrak{W})=0$ implies that $F_L(\mathfrak{W})=0$ for all $L\in\mathbb{N}$. Then Lemma 3 implies that

$$F_L(\mathfrak{W}_1\otimes\mathfrak{W}_2)=0 \tag{10}$$

as well which then gives us (7) as desired.

"$\Rightarrow$" Now we show the other direction. Assume that (7) holds. This implies that (10) is satisfied so that Lemma 3 yields (9). Thus, (8) is true proving the other direction. ∎

Remark 2. This result agrees with [11] where it has been shown for traditional decoding, i.e., $L=1$, that super-activation is a unique feature of secure communication over orthogonal AVCs and that it is not possible for public (non-secure) communication.

The next result shows that the list-L capacity is super-additive.

Theorem 6. Let $\mathfrak{W}_1$ and $\mathfrak{W}_2$ be two orthogonal AVCs. Then

$$C_L(\mathfrak{W}_1\otimes\mathfrak{W}_2)>C_L(\mathfrak{W}_1)+C_L(\mathfrak{W}_2) \tag{11}$$

if and only if

$$\min\big\{F_L(\mathfrak{W}_1),F_L(\mathfrak{W}_2)\big\}=0, \tag{12a}$$
$$\max\big\{F_L(\mathfrak{W}_1),F_L(\mathfrak{W}_2)\big\}>0, \tag{12b}$$

and

$$\min\big\{C_{CR}(\mathfrak{W}_1),C_{CR}(\mathfrak{W}_2)\big\}>0. \tag{13}$$

Proof: "$\Rightarrow$" If (12) is not true, then $C_L(\mathfrak{W}_i)=C_{\mathrm{CR}}(\mathfrak{W}_i)$, $i=1,2$. From Lemma 3 from the additivity of C_{CR} [11] we also know that

$$\begin{aligned}C_L(\mathfrak{W}_1\otimes\mathfrak{W}_2)&=C_{\mathrm{CR}}(\mathfrak{W}_1\otimes\mathfrak{W}_2)\\&=C_{\mathrm{CR}}(\mathfrak{W}_1)+C_{\mathrm{CR}}(\mathfrak{W}_2)\\&=C_L(\mathfrak{W}_1)+C_L(\mathfrak{W}_2)\end{aligned}$$

which contradicts (11) so that condition (12) must be satisfied.

If (13) is not true, then w.l.o.g. we have $C_{\mathrm{CR}}(\mathfrak{W}_1)=0$. We also have

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

From Theorem 5 we then know that either

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

or

$$C_L(\mathfrak{W}_1\otimes\mathfrak{W}_2)=0$$

which contradicts (11) so that condition (13) must be satisfied as well. This shows that both (12) and (13) are necessary for (11) to be true.

"$\Leftarrow$" Next, we show that (12) and (13) are also sufficient conditions. Assume (12) is true. Then w.l.o.g. we have $C_L(\mathfrak{W}_1)=0$ and from the additivity of C_{CR} [11]

$$\begin{aligned}C_L(\mathfrak{W}_1\otimes\mathfrak{W}_2)&=C_{\mathrm{CR}}(\mathfrak{W}_1\otimes\mathfrak{W}_2)\\&=C_{\mathrm{CR}}(\mathfrak{W}_1)+C_{\mathrm{CR}}(\mathfrak{W}_2)\\&>C_L(\mathfrak{W}_1)+C_L(\mathfrak{W}_2)\end{aligned}$$

which shows (11) proving the desired result. ∎

D. Bounds for Super-Additivity and the Problem of Alon

Alon studied how much the additivity of the zero error capacity can be violated. This question can be rewritten in terms of AVCs with maximum error.

In more detail, in [8] it was shown that for every $\epsilon > 0$ there exist two AVCs $\mathfrak{W}_1$ and $\mathfrak{W}_2$ with

$$C_{\text{max}}(\mathfrak{W}_i) \leq \epsilon, \quad i = 1, 2$$

such that for the normalized Shannon capacity

$$\bar{C}_{\text{max}}(\mathfrak{W}_i) = \frac{C_{\text{max}}(\mathfrak{W}_i)}{\log_2(\min(|\mathcal{X}_i|, |\mathcal{Y}_i|))}, \quad i = 1, 2$$

it holds that

$$\bar{C}_{\text{max}}(\mathfrak{W}_1 \otimes \mathfrak{W}_2) > \frac{1}{2} - \epsilon. \tag{14}$$

Remark 3. For the zero error capacity, Alon asked in [8] whether the constant $1/2$ in (14) can be replaced by 1. This question was negatively answered in [9] by showing that the constant $1/2$ is indeed tight in this case.

In the following, we want to study this question for AVCs under list decoding. Let $\mathfrak{W}_1$ and $\mathfrak{W}_2$ be two arbitrary finite AVCs with

$$C_L(\mathfrak{W}_i) \leq \epsilon, \quad i = 1, 2. \tag{15}$$

If $\max\{F_L(\mathfrak{W}_1), F_L(\mathfrak{W}_2)\} = 0$, then it follows from Theorem 5 that

$$C_L(\mathfrak{W}_1 \otimes \mathfrak{W}_2) = 0$$

so that there is nothing to prove. If we have $\min\{F_L(\mathfrak{W}_1), F_L(\mathfrak{W}_2)\} > 0$, then the list capacity C_L is additive and there is also nothing to prove. Therefore, it remains to study the case for which (15) holds and further

$$\min\{F_L(\mathfrak{W}_1), F_L(\mathfrak{W}_2)\} = 0.$$

W.l.o.g. we assume that $F_L(\mathfrak{W}_1) = 0$ so that we must have $C_{\text{CR}}(\mathfrak{W}_1) > 0$ for super-additivity. Thus, we have

$$\begin{aligned} C_L(\mathfrak{W}_1 \otimes \mathfrak{W}_2) &= C_{\text{CR}}(\mathfrak{W}_1) + C_L(\mathfrak{W}_2) \\ &\leq C_{\text{CR}}(\mathfrak{W}_1) + \epsilon. \end{aligned}$$

For sufficiently small ϵ, we choose a DMC W_2 with $C(W_2) = \epsilon$ which is always possible. With $\mathfrak{W}_2 = \{W_2\}$ we have $C_L(\mathfrak{W}_2) = C_L(W_2) = \epsilon$ and therewith

$$\sup_{\mathfrak{W}_1, \mathfrak{W}_2} C_L(\mathfrak{W}_1 \otimes \mathfrak{W}_2) = \sup_{\mathfrak{W}_1} C_{\text{CR}}(\mathfrak{W}_1) + \epsilon$$

where the sup is over all AVCs that satisfy (15) where $\mathfrak{W}_1$ must further satisfy $F_L(\mathfrak{W}_1) = 0$ and for all $\epsilon > 0$ there is a finite AVC $\hat{\mathfrak{W}}$ with with $D(\hat{\mathfrak{W}}, \mathfrak{W}_1) < \epsilon$ and $F_L(\hat{\mathfrak{W}}) > 0$. Next, we want to construct a good upper bound on the normalized list-L capacity

$$\bar{C}_L(\mathfrak{W}) = \frac{C_L(\mathfrak{W})}{\log_2(\min(|\mathcal{X}|, |\mathcal{Y}|))}. \tag{16}$$

For sufficiently small ϵ, we can choose $|\mathcal{X}_2| = |\mathcal{Y}_2| = 2$. With this choice, we obtain

$$\bar{C}_L(\mathfrak{W}_1 \otimes \mathfrak{W}_2) \leq \frac{C_{\text{CR}}(\mathfrak{W}_1) + \epsilon}{\log_2(\min(|\mathcal{X}_1|, |\mathcal{Y}_1|)) + 1}.$$

From [5, Theorem 1] we know that $C_{\text{CR}}(\mathfrak{W}_1) \leq \frac{1}{L}\log_2(\min(|\mathcal{Y}_1|, |\mathcal{S}_1|))$ so that we obtain for the normalized list-L capacity

$$\begin{aligned} \bar{C}_L(\mathfrak{W}_1 \otimes \mathfrak{W}_2) \leq \frac{1}{L}\Big(\frac{\log_2(\min(|\mathcal{Y}_1|, |\mathcal{S}_1|))}{\log_2(\min(|\mathcal{X}_1|, |\mathcal{Y}_1|)) + 1}\Big) \\ + \frac{\epsilon}{\log_2(\min(|\mathcal{X}_1|, |\mathcal{Y}_1|)) + 1}. \end{aligned}$$

This means that the normalized list-L capacity approaches zero as $L \to \infty$. In other words, the normalized list-L-capacity (16) becomes smaller with increasing list size L. This is in contrast to the behavior we observe in (14) for the zero error capacity. We further see that the strong form of violation does not hold; cf. also Remark 3.

IV. Conclusion

A unified theory has been developed that enables the study of properties of capacity functions such as discontinuity and non-additivity. In particular, in this paper the list-L capacity of an AVC has been studied and we have shown that it is a discontinuous function. In particular, the discontinuity points have been completely characterized. We have further shown that the list-L capacity function is super-additive and have characterized those AVCs that possess this property. As a consequence, joint encoding and decoding can result in a higher list-L capacity than processing every orthogonal AVC individually. While this behavior already appears for the deterministic capacity C (i.e. list size $L = 1$), we have shown that the discrepancy for the list-L capacity can further be arbitrarily large. The gain in list-L capacity due to joint processing increases without bound with increasing L.

References

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

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

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

[4] V. Blinovsky, P. Narayan, and M. Pinsker, "Capacity of the arbitrarily varying channel under list decoding," *Probl. Pered. Inform.*, vol. 31, no. 2, pp. 99–113, 1995.

[5] B. L. Hughes, "The smallest list for the arbitrarily varying channel," *IEEE Trans. Inf. Theory*, vol. 43, no. 3, pp. 803–815, May 1997.

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

[7] W. Haemers, "On some problems of Lovász concerning the Shannon capacity of a graph," *IEEE Trans. Inf. Theory*, vol. 25, no. 2, pp. 231–232, Mar. 1979.

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

[9] P. Keevash and E. Long, "On the normalized Shannon capacity of a union," *Comb. Probab. Comp.*, vol. 25, no. 5, pp. 766–767, Sep. 2016.

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

[11] R. F. Schaefer, H. Boche, and H. V. Poor, "Super-activation as a unique feature of arbitrarily varying wiretap channels," in *Proc. IEEE Int. Symp. Inf. Theory*, Barcelona, Spain, Jul. 2016, pp. 3077–3081.