Symmetrizability for Myopic AVCs

Amitalok J. Budkuley*, Bikash Kumar Dey†, Sidharth Jaggi‡,
Michael Langberg§, Anand D. Sarwate¶, Carol Wang
*IIT Kharagpur, †IIT Bombay, ‡CUHK, §SUNY Buffalo, ¶Rutgers University

***Abstract*—Myopic arbitrarily varying channels (AVCs) are point-to-point communication models in which a channel state is controlled by a malicious adversary (a jammer) who receives side-information about the transmitted codeword via a side-channel (wiretapping) and wishes to maximize the probability of error. Compared to standard "oblivious" AVCs, myopic AVCs can potentially use the side information to launch a more effective attack, lowering the capacity of the channel. In this paper, we define a novel property, myopic symmetrizability, and prove it is a sufficient condition for the capacity of any myopic AVC to be zero. We also study the *sufficiently myopic* setting, in which, roughly speaking, the jammer's side information reveals less information on the codeword transmitted than eventually available at the receiver. In this scenario we show that myopic symmetrizability is also a necessary condition for the capacity to equal zero, by providing a novel code construction using non-i.i.d. codebooks. A key technical lemma, interesting in its own right, is an argument showing that for *any* positive-rate code (whether for myopic AVCs or not) one can identify a corresponding distribution $P_{X,X'}$ that is a convex combination of product distributions, and such that a constant fraction of pairs of codewords have an empirical distribution approximately equaling $P_{X,X'}$.**

I. Introduction

The arbitrarily varying channel (AVC) is a model for communication over a channel whose time-varying state is controlled by an adversary (cf. [1], [2]). In keeping with prior convention we call the encoder Alice, the decoder Bob and the adversary, or jammer, James. In the case where the input codeword $\mathbf{x}$ and state sequence $\mathbf{s}$ may be constrained (e.g. by total per-letter costs), the capacity under the average probability of error criterion was characterized by Csiszár and Narayan [3], who showed that if James could *symmetrize* the channel, then the capacity is zero. Loosely speaking, James can symmetrize the channel if he can "spoof" a valid codeword: he can make the channel "look like" a symmetric two-user (Alice and James) multiaccess channel. James randomly chooses a codeword $\mathbf{x}'$ to send and Bob will not be able to disambiguate reliably between Alice's $\mathbf{x}$ and James's $\mathbf{x}'$.

Recent works have proposed modifications of the general AVC to investigate how James's knowledge of the transmitted codeword affects the capacity problem. In this paper we study one such variant of this model, the *myopic AVC*, in which the adversary can observe the transmitted channel codeword through a discrete memoryless channel (DMC) before choosing a state sequence to maximize the probability of error [4]–[6]. The myopic AVC has some similarities to wiretap models, although here the adversary can affect the channel state with the goal of maximizing the probability of error. Dey, Jaggi, and Langberg [6] showed that for "sufficiently myopic" adversaries[1], the capacity is the same as if the adversary had no knowledge of the transmitted codeword.

The notion of symmetrization for myopic AVCs is a more complex phenomenon than for the commonly studied model of AVCs [3], referred to here as *oblivious* AVCs (in the sense that the jammer has no knowledge of the transmitted codeword $\mathbf{x}$). In symmetrizability for oblivious AVCs, James may choose any valid codeword $\mathbf{x}'$ for his attack: that is, he can render any pair of codewords $(\mathbf{x}, \mathbf{x}')$ confusable. In the myopic case, specific pairs of codewords may not be confusable but for any sufficiently large codebook there must exist many pairs which are confusable. In addition, James may tailor his jamming attack on the basis of his noisy observation of $\mathbf{x}$.

The primary focus of this work is on understanding whether a positive rate is possible or not for a given myopic AVC. To this end our main contribution is to define the new notion of *myopic symmetrizability* that acts as a *sufficient* condition for a myopic AVC to admit no positive rate. Namely, we show that the capacity is zero for myopically symmetrizable AVCs. For the interesting subclass of sufficiently myopic AVCs [6] we show that this is also a *necessary* condition, by giving lower (achievable) bounds on the capacity for myopically non-symmetrizable AVCs.

A key technical lemma in our zero-rate converse, interesting in its own right, is an argument showing that for *any* positive-rate code (whether for myopic AVCs or not) one can identify a corresponding distribution $P_{X,X'}$ over $\mathcal{X} \times \mathcal{X}$ such that it can be decomposed as a convex combination of product distribution, and such that at least a constant fraction of pairs of codewords have an empirical distribution approximately equaling $P_{X,X'}$. The guarantee afforded by this lemma implies that if the AVC is myopically symmetrizable, even if the malicious adversary only attempts to confuse the receiver about such pairs of codewords, he will nonetheless succeed with constant probability. The following example, of a myopic bit-flip AVC, previews the flavour of our main results.

Example 1. *A "myopic (p, q) bit-flip AVC" may be described as follows: The transmitter Alice wishes to communicate a message m to the receiver Bob over n uses of a binary-input binary-output channel by transmitting a length-n binary codeword $\mathbf{x}$. Alice's encoder is known a priori to all parties. The malicious jammer James observes $\mathbf{z}$, a noisy version of $\mathbf{x}$ corresponding to passing $\mathbf{x}$ through a binary symmetric*

[1]Roughly speaking, a sufficiently myopic AVC is one in which the worst DMC James can induce from Alice to Bob has a higher Shannon capacity than the channel from Alice to James.

channel (BSC) with a transition probability q, i.e., a BSC(q). As a function of his observation $\mathbf{z}$ and his knowledge of Alice's encoder, James may choose an arbitrary jamming sequence $\mathbf{s}$ of Hamming weight at most pn. Bob receives $\mathbf{y} = \mathbf{x} \oplus \mathbf{s}$ (where $\oplus$ denotes component-wise modulo-2 addition). From $\mathbf{y}$ and his knowledge of Alice's encoder Bob attempts to recover her message m.

For this AVC, as a consequence of Theorem 4 (our first main result, stated in Section III) we can show that for $0 \leq q < 1/2$ no positive rate is possible if $p \geq \frac{1}{4(1-q)}$. This novel impossibility result comes from interpolating between known impossibility results when $q = 1/2$ (the oblivious setting [3] – in this case communication is impossible if and only if $p \geq 1/2$), and $q = 0$ (the omniscient setting – in this case the classical Plotkin bound [7] implies no code of size larger than $\mathcal{O}(1/\varepsilon)$ exists if $p \geq 1/4 + \varepsilon$). Our impossibility result for this example may be understood as appropriately fusing the probabilistic arguments for oblivious channels [3] and the combinatorial arguments for omniscient channels [7].[2] *For general AVCs, Theorem 4 similarly combines oblivious symmetrizability arguments [3] with the recently developed Generalized Plotkin Bound [8].*

Our achievability result in Theorem 7 (our second main result, stated in Section III) is restricted to the sufficiently myopic setting, which in this example corresponds to p and q such $1 - H(p) < 1 - H(q)$, i.e. $p < q < 1/2$. Using the notation outlined in Sections II and III, in this parameter regime the optimizing $P_{U,X}$ distribution in Theorem 7 collapses just to the Bernoulli($1/2$) i.i.d. distribution on X with no need of the time-sharing variable U, and the corresponding achievable rate guaranteed by the theorem equals $1 - H(p)$. In this regard Theorem 7 retrieves a prior result [6]. However, it is possible to construct AVCs for which i.i.d. distributions may not achieve a positive rate, whereas independent but not identically distributed codebooks can – Example 1 in Wang et al. [8] demonstrates such an (omniscient) AVC.[3]

It was noted in [9] that stochastic code constructions (with private randomness at the encoder unknown *a priori* to either the jammer or the decoder) can outperform deterministic code constructions for *insufficiently myopic* AVCs.[4] This is one reason why in this work we do not settle the rate-positivity question for insufficiently myopic AVCs, since here for our inner bounds we consider only deterministic code constructions (though our myopic symmetrizability impossibility result holds in great generality, even for stochastic codes). In ongoing work we are investigating extending the deterministic code constructions in this paper to stochastic codes, and examining whether this can help complete the characterization of the rate-positivity question for all myopic AVCs, whether sufficiently myopic or not.

We note that the current results reported here correspond to understanding when the rate is positive: our achievable rates may not be capacity-achieving. Understanding the optimal rate for merely sufficiently myopic AVCs is an interesting problem in its own right that we do not address. However, we believe that a tight rate-converse may be possible. Indeed, proving a rate-converse that would be tight in this regime can follow from a combinatorial conjecture (stated at the end of the paper) that may be interesting in its own right.

II. Channel Model

Notation: Vectors or sequences will be given in boldface and sets in calligraphic script. For a positive integer a the set $[a]$ denotes $\{1, 2, \ldots, a\}$. Let $\Delta(\mathcal{X})$ be the set of all probability mass functions on a finite alphabet $\mathcal{X}$ (i.e. the simplex of probability distributions). For distributions Q and Q' on $\mathcal{X}$, let $d_\infty(Q, Q') = \|Q - Q'\|_\infty$ and define the ℓ_∞ ball $B_\infty(Q, \delta) = \{Q' \in \Delta(\mathcal{S}) : \|Q - Q'\|_\infty < \delta\}$. For a set $\Lambda \subseteq \Delta(\mathcal{X})$ define the δ-interior $\mathrm{Int}_\delta(\Lambda) = \{Q \in \Lambda : B_\infty(Q, \delta) \subseteq \Lambda\}$ and the distance $d_\infty(Q, \Lambda) = \inf_{Q' \in \Lambda} d_\infty(Q, Q')$. For an n-tuple $\mathbf{x} \in \mathcal{X}^n$, the type (empirical distribution) of $\mathbf{x}$ is denoted by $T_{\mathbf{x}} \in \Delta(\mathcal{X})$. For a joint distribution $P_{X_1 X_2 \cdots X_m}$ on variables $X_1, X_2, \ldots, X_m$, we denote the marginal distribution on X_i by $[P_{X_1 X_2 \cdots X_m}]_{X_i}$.

Let $\mathcal{X}$, $\mathcal{Z}$, $\mathcal{S}$, and $\mathcal{Y}$ be finite alphabets. A *myopic arbitrarily varying channel* $\mathcal{A}$ consists of a discrete memoryless channel (DMC) $W_{Z|X}(z|x)$ from $\mathcal{X}$ to $\mathcal{Z}$ and a collection $\{W_{Y|X,S}(y|x,s) : s \in \mathcal{S}\}$ of DMCs indexed by $\mathcal{S}$. For a fixed $\mathbf{x} \in \mathcal{X}^n$ and $\mathbf{s} \in \mathcal{S}^n$ the distributions of $\mathbf{z}$ and $\mathbf{y}$ are given by $W^n_{Z|X}(\mathbf{z}|\mathbf{x}) = \prod_{i=1}^n W_{Z|X}(z_i|x_i)$, and $W^n_{Y|X,S}(\mathbf{y}|\mathbf{x},\mathbf{s}) = \prod_{i=1}^n W_{Y|X,S}(y_i|x_i,s_i)$. If Z is independent of X then this reduces to the standard AVC, which we call an *oblivious AVC* and if $Z = X$ we call it an *omniscient AVC*. A (n, N) *code with stochastic encoding* for a myopic AVC is a pair of maps (Φ, ψ) where $\Phi : [N] \to \mathcal{X}^n$ is a stochastic encoding map and $\psi : \mathcal{Y}^n \to [N]$ is a deterministic decoding map. A (n, N) *deterministic code* is a special case in which Φ is replaced with a deterministic mapping $\phi : [N] \to \mathcal{X}^n$. The rate of an (n, N) code is $R = \frac{1}{n} \log_2(N)$.

In an AVC we assume the state $\mathbf{s}$ is controlled by a malicious adversary, or jammer, who wishes to limit the maximum rate of reliable communication. In a myopic AVC the jammer has access to side information about the transmitted codeword. If $m \in [N]$ is to be sent and $\mathbf{x} = \phi(m)$, the jammer can observe $\mathbf{z}$, a noisy version of $\mathbf{x}$. We assume that the jammer knows the code being used by the encoder/decoder but not the message being transmitted. A jamming strategy is a map $\mathcal{J} : \mathcal{Z}^n \to \mathcal{S}^n$. That is, the jammer can see the entire sequence $\mathbf{z} \in \mathcal{Z}^n$ and uses that to choose its input $\mathbf{s} \in \mathcal{S}^n$.

[2] This parameter regime $p \geq \frac{1}{4(1-q)}$ corresponds solely to insufficiently myopic AVCs – however, with some care one can even construct sufficiently myopic AVCs which are nonetheless myopically symmetrizable.

[3] While Example 1 in Wang et al. [8] is *insufficiently myopic* since it corresponds to an omniscient AVC, it can be bootstrapped into a more complex example which is *sufficiently myopic*, and still exhibits similar behaviour, where no i.i.d. codebook achieves positive rate, but there exists a non-i.i.d. distribution resulting in a code that achieves positive rate.

[4] An example of a myopic AVC was provided in Dey et al.! [9] demonstrating not only that stochastic codes could achieve higher rates than deterministic codes, but also that the parameter regimes where positive rates could be attained was strictly larger with stochastic codes than with deterministic codes.

For a code (Φ, ψ), message m, and strategy $\mathcal{J}$ the probability of error is

$$\varepsilon(m, \mathcal{J}) = \sum_{\mathbf{x},\mathbf{z},\mathbf{s}} W^n_{Y|X,S}(\psi^{-1}(m)|\mathbf{x}, \mathbf{s}) W^n_{Z|X}(\mathbf{z}|\mathbf{x}) \mathbb{P}(\Phi(m) = \mathbf{x}) \mathbb{P}(\mathcal{J}(\mathbf{z}) = \mathbf{s}). \quad (1)$$

The (average) probability of error is defined as $\bar{\varepsilon} = \max_{\mathcal{J}} \frac{1}{N} \sum_{m=1}^{N} \varepsilon(m, \mathcal{J})$, where the maximum is over $\mathcal{J}$ satisfying the cost constraints as defined below. A rate R is achievable (under average error) if for any $\epsilon > 0$ there exists a sequence of $(n, 2^{nR})$ codes such that $\bar{\varepsilon} < \epsilon$. The capacity is the supremum of achievable rates over all codes satisfying the constraints below.

We define linear constraints on codes and jamming strategies as follows. Let $\Lambda_s \subseteq \Delta(\mathcal{S})$ and $\Lambda_x \subseteq \Delta(\mathcal{X})$ be convex polytopes. A code satisfies the constraint Λ_x if the codeword type $T_{\Phi(m)} \in \Lambda_x$ for all m almost surely. A jamming strategy $\mathcal{J}$ satisfies the constraint Λ_s if the type $T_{\mathbf{s}} \in \Lambda_s$ almost surely. As an example suppose $\mathcal{X} = \mathcal{S} = \{0, 1\}$ and Λ_x are all distributions P with $P(1) \leq \alpha$ and Λ_s contains all distributions Q with $Q(1) \leq \beta$. Then, almost surely, codewords have Hamming weight at most αn and jamming sequences have Hamming weight at most βn.

III. Main results

In this work we generalize the symmetrizability property for oblivious AVCs given by Csiszár and Narayan [3]. To make the contrast with our new notion we call their definition *oblivious* symmetrizability. Due to space constraints, the proofs of all claims below will appear in the extended version of this work.

Definition 1 (Oblivious symmetrizability [3])**.** *For an oblivious AVC $\mathcal{A}$, an input distribution P_X is obliviously symmetrizable if there exists a channel $V_{S|X'}$ from $\mathcal{X}$ to $\mathcal{S}$ such that the following two conditions hold:*

(i) For all $(x, x', y) \in \mathcal{X} \times \mathcal{X} \times \mathcal{Y}$:

$$\sum_s W_{Y|X,S}(y|x, s) V_{S|X'}(s|x') = \sum_{s'} W_{Y|X,S}(y|x', s') V_{S|X'}(s'|x), \quad (2)$$

(ii) $\exists\ \delta > 0$ such that $[P_X V_{S|X}]_S \in \mathrm{Int}_\delta(\Lambda_s)$.

As mentioned in the introduction, the condition in (2) means an oblivious adversary can use the channel $V_{S|X'}$ to create a symmetric multiaccess channel with inputs X and X' such that the distribution of the output $\mathbf{y}$ is the same with inputs $(\mathbf{x}, \mathbf{x}')$ or $(\mathbf{x}', \mathbf{x})$.

To define myopic symmetrizability, we will use the following notion of *completely positive* (CP) joint distributions.

Definition 2 (Completely Positive (CP) distributions)**.** *For a distribution P_X on $\mathcal{X}$, the set of* completely positive joint distributions $\mathsf{CP}(P_X)$ *is defined as the set of joint distributions $P_{X,X'} \in \Delta(\mathcal{X} \times \mathcal{X})$ such that*

(i) The marginal distributions are equal to P_X: $[P_{X,X'}]_X = [P_{X,X'}]_{X'} = P_X$.

(ii) There exists a finite alphabet $\mathcal{U}$ of cardinality at most $|\mathcal{X}|(|\mathcal{X}| - 1)/2$ and distribution P_U on $\mathcal{U}$ such that $P_{X,X'}(x, x') = \sum_{u \in \mathcal{U}} P_U(u) P_{X|U}(x|u) P_{X|U}(x'|u)$.

For a subset of distributions $\Gamma \subset \Delta(\mathcal{X})$ the set $\mathsf{CP}(\Gamma) = \bigcup_{P \in \Gamma} \mathsf{CP}(P)$.

A completely positive distribution $P_{X,X'}$ may be viewed as the expected joint type between pairs of codewords in a code ensemble where for each $u \in \mathcal{U}$, in expectation $nP_U(u)$ codeword coordinates are sampled according to the distribution $P_{X|U}(\cdot|U = u)$.[5]

Definition 3 (Myopic symmetrizability)**.** *For a myopic AVC $\mathcal{A}$, a CP distribution $P_{X,X'}$ (and the corresponding P_{UX}) is* myopically symmetrizable *if there exists a channel $V_{S|Z,X'}$ such that the following two conditions hold:*

(i) For all $(x, x', y) \in \mathcal{X} \times \mathcal{X} \times \mathcal{Y}$:

$$\sum_{z,s} W_{Z|X}(z|x) V_{S|Z,X'}(s|z, x') W_{Y|X,S}(y|x, s) = \sum_{z',s'} W_{Z|X}(z'|x') V_{S|Z,X'}(s'|z', x) W_{Y|X,S}(y|x', s'), \quad (3)$$

(ii) $\exists\ \delta > 0$ such that $[P_{X,X'} W_{Z|X} V_{S|ZX'}]_S \in \mathrm{Int}_\delta(\Lambda_s)$.

If no such channel exists we call the distribution myopically non-symmetrizable.

For a given AVC $(W_{Z|X}, W_{Y|X,S})$, we denote the set of myopically non-symmetrizable distributions $P_{X,X'}$ as $\mathcal{NS}$. Our first result shows that myopic symmetrizability is a sufficient condition for zero capacity.

Theorem 4. *A positive rate is not possible for a myopic AVC $\mathcal{A}$ if the set $\mathcal{NS}$ of myopically non-symmetrizable distributions is empty.*

A few remarks are in order. Definition 1 differs from Definition 3 in two major aspects. First, as one would expect, in condition (i) of Definition 3 James's view z of the transmitted symbol x through $W_{Z|X}$ plays a central role. The symmetrizing mapping $V_{S|Z,X'}$ is thus a function of both x' and z. For readers familiar with the oblivious AVC techniques [3] such a mapping is the natural generalization. Secondly and more subtly, in condition (ii) of Definition 3, we consider the joint distribution $P_{X,X'} \in \mathsf{CP}(P_X)$ and not just P_X as in Definition 1. Such a restriction of $P_{X,X'}$ is a non-trivial strengthening. For instance, if we did not require $P_{X,X'}$ to be a CP distribution in Example 1, the symmetrizing region would collapse to the oblivious symmetrizing region $p \geq 1/2$, rather than the strictly larger region $p \geq \frac{1}{4(1-q)}$ we obtain with the restriction. This is because if $P_{X,X'}$ is not

[5]The set of *completely positive matrices* are known (see for instance the excellent survey by Berman and Shaked-Monderer [10]) to form a closed convex cone. Our definition of $\mathsf{CP}(P_X)$ may be viewed as intersecting this cone with additional hyperplanes – in particular requiring the row/column sums to equal P_X. The cardinality bound of $|\mathcal{X}|(|\mathcal{X}|-1)/2$ arises by noting that $\mathsf{CP}(P_X)$ is a convex set comprising of $|\mathcal{X}| \times |\mathcal{X}|$ symmetric matrices with row/column sums both pre-specified as P_X, and applying Carathéodory's extension theorem.

required to be a CP distribution, then James would also have to symmetrize distributions in which $x \neq x'$ with probability 1, which requires the Hamming weight of $\mathbf{s}$ to equal $n/2$ even if James is omniscient.

The proof of Theorem 4 relies heavily on the Generalized Plotkin Bound proven by Wang et al. [8]. In particular, we need the following "robust" version of the converse argument in Theorem 9 of that work.

Lemma 5 (Robust Generalized Plotkin Bound)**.** *Let Λ_x be a constraint set for codewords over $\mathcal{X}$, n be a positive integer, and $\mathcal{C}$ be a collection of sequences in $\mathcal{X}^n$ (a codebook) such that $T_{\mathbf{x}} \in \Lambda_x$ for all $\mathbf{x} \in \mathcal{C}$. Then for any $\delta > 0$, if $d_\infty(T_{\mathbf{x}\mathbf{x}'}, \mathsf{CP}(\Lambda_x)) > \delta$ for all $(\mathbf{x}, \mathbf{x}') \in \mathcal{C} \times \mathcal{C}$, then $|\mathcal{C}| \leq \mathcal{O}(1/\delta^{2|\mathcal{X}|^2})$.*

Proof sketch: We highlight, for deterministic codes, the intuition behind the proof of Theorem 4. The proof can be generalized to the setting of stochastic codes using the line of arguments presented in prior work [11].

Let $\mathcal{C}$ be any code of positive rate shared by Alice and Bob. We first claim that there exists a large *approximately-constant composition* subcode $\mathcal{C}'$ of $\mathcal{C}$. Namely, that there exists a subcode $\mathcal{C}' \subseteq \mathcal{C}$ of size $\Omega(|\mathcal{C}|)$ for which each codeword has approximate type $\hat{P}_X$. This follows from considering, for constant $\delta > 0$, a δ-net quantization of the simplex $\Delta(\mathcal{X})$ and the corresponding Voronoi partition. As $\mathcal{X}$ is finite, such a net is also finite, and induces a finite partition on $\mathcal{C}$. The largest subcode $\mathcal{C}'$ in this partition, obtained by the Voronoi cell with center $\hat{P}_X$, will have size $\Omega(|\mathcal{C}|)$.

We now consider pairs of codewords in $\mathcal{C}'$ and claim that a constant fraction of them are δ-close to some $\hat{P}_{X,X'} \in CP(\Lambda_X)$. This follows from a combination of Lemma 5, Turán's Theorem, and the fact that $CP(\Lambda_X)$ can be quantized as well into a finite sized δ-net. Specifically, by Lemma 5, there do not exist large subsets of $\mathcal{C}'$ for which all pairs of codewords are far from $CP(\Lambda_X)$. Turán's Theorem then implies that a constant fraction of pairs of codewords in $\mathcal{C}'$ must be δ-close to $CP(\Lambda_X)$. This, in turn, combined with a δ-net on $CP(\Lambda_X)$ implies our claim.

Consider now a uniform randomly chosen codeword $\mathbf{x}$ from $\mathcal{C}$. As $\hat{P}_X$ is myopically symmetrizable, James may select the channel $V_{S|ZX'}$ affiliated with $\hat{P}_{X,X'}$, and a spoofing codeword $\mathbf{x}'$ uniformly and random from the codebook $\mathcal{C}$. By the discussion above, with constant probability the joint type of $\mathbf{x}, \mathbf{x}'$ is δ-close to $\hat{P}_{X,X'}$. James then selects his jamming vector $\mathbf{s}$ by applying $V_{S|Z,X'}$ symbol-wise on $\mathbf{z}$ and $\mathbf{x}'$. By our definition of symmetrizability, Bob will not be able to distinguish between the case at hand, and that in which Alice transmitted $\mathbf{x}'$. ∎

A. The sufficiently-myopic case

Our second main result addresses coding schemes for AVCs that satisfy a generalized version of sufficiently myopic AVCs defined in [6]. Here, we address only non-symmetrizable $P(X)$ for which by Definition 1 there exist $P_{X,X'} \in \mathsf{CP}(P_X)$ with certain properties. We thus use the notation $P_{U,X}$ to describe such P_X that arise from the CP-decomposition given in Definition 2.

Definition 6 (Sufficiently myopic AVCs)**.** *We say that an AVC is sufficiently myopic under $P_{U,X}$ if the quantity*

$$R^*(P_{U,X}, W_{Z|X}, W_{Y|X,S}) = \min_{V_{S|Z,U}} I(X;Y|U)$$

is strictly larger than $I(X;Z|U)$, i.e., if the worst DMC James can induce from Alice to Bob has a higher Shannon capacity than the channel from Alice to James.

We define

$$\mathcal{P} := \{P_{U,X} | \text{the AVC is sufficiently myopic under } P_{U,X}\}$$

In Theorem 7 below, $\mathcal{NS}$ denotes the set of $P_{U,X}$ that are non-symmetrizable.

Theorem 7. *If $\mathcal{P}$ is non-empty for an AVC, then the rate*

$$\max_{P_{U,X} \in \mathcal{NS} \cap \mathcal{P}} \min_{V_{S|Z,U}} I(X;Y|U).$$

is always achievable. Here U is a (time-sharing) random variable with a cardinality bound[6] given by $|\mathcal{U}| \leq |\mathcal{X}|(|\mathcal{X}|-1)/2$.

Proof sketch: Key to the construction of the code is a distribution $P_{U,X}$ corresponding to a completely positive distribution $P_{X,X'}$. Given such a distribution we can identify a random variable U on an alphabet $\mathcal{U}$ with distribution P_U such that $P_{X,X'} = [P_U P_{X|U} P_{X'|U}]_{X,X'}$, where the conditional distributions $P_{X|U}$ and $P_{X'|U}$ are the same. Let us fix the pair $(P_U, P_{X|U})$. We construct the codebook as follows:

1) Fix any $\mathbf{u} \in \mathcal{U}^n$ with type $T_{\mathbf{u}} = P_U$. We call $\mathbf{u}$ the cloud centre.
2) For $m \in [2^{nR}]$ generate the codeword $\mathbf{x}(m)$ according to the distribution $\prod_{i=1}^n P_{X|U}(x_i|u_i)$.

The cloud centre $\mathbf{u}$ and codebook $\mathcal{C} = \{\mathbf{x}(m) : m \in [2^{nR}]\}$ are revealed to Alice, James, and Bob. To encode a message m, the encoder sends $\mathbf{x}(m)$ if $d_\infty(T_{\mathbf{x}(m)}, P_X) < \delta$ and $T_{\mathbf{x}(m)} \in \Lambda_x$, otherwise it declares an error.

The corresponding decoding scheme and error-analysis is broadly similar to that used in oblivious AVCs [3]. To decode, Bob uses a two-step decoder.

1. List decoding: Bob first determines a list $\mathcal{L}(\mathbf{y})$ of candidate codewords $\mathbf{x} \in \mathcal{C}$ such that for each codeword $\mathbf{x}$ in the list there exist $\mathbf{z}$, feasible $\mathbf{s}$ and a channel $V_{S|Z,U}$ such that

$$d_\infty(T_{\mathbf{uxzsy}}, P_U P_{X|U} W_{Z|X} V_{S|Z,U} W_{Y|X,S}) \leq \eta,$$
$$[P_U P_{X|U} W_{Z|X} V_{S|Z,U}]_S \in \Lambda_s$$

where $\eta > 0$ is a constant. Relatively standard techniques in the AVC literature ensure that for sufficiently myopic AVCs (this is the only place where sufficient myopia is needed) with high probability over code-design, the cardinality of Bob's list is at most polynomial in n.

[6]We remark on a subtle point here – while $|\mathcal{X}|(|\mathcal{X}|-1)/2$ suffices as a cardinality bound for a CP-decomposition of a given $P_{X,X'}$ over an alphabet $\mathcal{X}$, there are always "redundant" decompositions with (arbitrarily) larger cardinality. It is not obvious whether or not the functional $I(X;Y|U)$ may have a larger optimizing value under some such redundant decomposition.

2. Tournament: We say $\mathbf{x}$ dominates $\mathbf{x}'$ if there exist $\mathbf{z}$, feasible $\mathbf{s}$ and $V_{S|Z,X',U}$ such that

$$d_\infty(T_{\mathbf{uxx'zsy}}, P_U P_{X|U} P_{X'|U} W_{Z|X} V_{S|Z,X',U} W_{Y|X,S}) \leq \eta,$$
$$[P_U P_{X|U} P_{X'|U} W_{Z|X} V_{S|Z,X',U}]_S \in \Lambda_s,$$

and for every $\mathbf{z}'$, feasible $\mathbf{s}'$ we have

$$d_\infty(T_{\mathbf{ux'xz's'y}}, P_U P_{X|U} P_{X'|U} W_{Z|X} V_{S|Z,X',U} W_{Y|X,S}) > \eta.$$

If there is a $\tilde{\mathbf{x}}$ that dominates all other $\mathbf{x}' \in \mathcal{L}(\mathbf{y})$ then Bob outputs the message corresponding to that $\tilde{\mathbf{x}}$. Otherwise he declares an error. Arguments similar to the oblivious case [3], [12] may be used to argue that due to myopic non-symmetrizability, with high probability the correct codeword $\mathbf{x}$ will win the tournament stage of decoding. ∎

The code construction in Theorem 7 results in deterministic codes, but as already noted in the Introduction, it is known [9] that at least for some AVCs stochastic codes strictly outperform *any* deterministic codes. We are therefore currently exploring in ongoing work the impact of incorporating encoder randomness via superposition coding, for instance in the classical code construction for broadcast [13] in the context of *passive* security against eavesdroppers, and the concomitant notions of stochastic myopic symmetrizability required.

IV. Discussion

In this paper we provide a notion of myopic symmetrizability which is a sufficient condition for the capacity of a myopic AVC to be zero, and affords a tight characterization of the positive-rate region for sufficiently myopic AVCs. However, we do not provide a tight rate-converse on the achievable rate once the region $\mathcal{NS}$ is non-empty. Towards this end, we pose an open problem which we find intriguing and which can lead to a rate-converse that would be tight in the sufficiently myopic regime. The statement of our open problem appears below, followed by a short discussion.

Let $\delta > 0$. Consider a given code $\mathcal{C}$. Let $P_{X,X'} \in \Delta(\mathcal{X} \times \mathcal{X})$ be defined by probability distributions $P_{X|U}$ and P_U through $P_{X,X'}(x, x') = \sum_{u \in \mathcal{U}} P_U(u) P_{X|U}(x|u) P_{X|U}(x'|u)$.

Open question: *If for a constant fraction of pairs of codewords $(\mathbf{x}, \mathbf{x}')$ in $\mathcal{C} \times \mathcal{C}$ it holds that $d_\infty(T_{\mathbf{x},\mathbf{x}'}, P_{X,X'}) \leq \delta$, then there exists a constant-fraction sized subcode $\mathcal{C}'$ of $\mathcal{C}$ and a single vector $\mathbf{u}$ such that for all $\mathbf{x} \in \mathcal{C}'$ it holds that $d_\infty(T_{\mathbf{x}|\mathbf{u}}, P_{X|U}) \leq \delta$.*

In words, the question above seeks to find a witness $\mathbf{u}$ to the structure of the code $\mathcal{C}$ in the sense that we assume (a constant fraction of) pairs of codewords in $\mathcal{C}$ have (approximate) joint type governed by the distribution P_U and we wish to prove the existence of a realization $\mathbf{u}$ that governs the (approximate) type of (a constant fraction of) codewords in $\mathcal{C}$.

The open question plays a role in a potential rate-converse using the following line of analysis. First, using standard averaging arguments any code $\mathcal{C}$ has an associated distribution $P_{X,X'}$ for which, for small $\delta > 0$, a constant fraction of pairs of codewords $(\mathbf{x}, \mathbf{x}')$ in $\mathcal{C} \times \mathcal{C}$ satisfy $d_\infty(T_{\mathbf{x},\mathbf{x}'}, P_{X,X'}) \leq \delta$. Using this observation with distributions that are completely positive, an affirmative answer to the open question above yields a witness $\mathbf{u}$ governing the type of codewords in a constant-fraction sized subcode of $\mathcal{C}'$. Now, carefully using $\mathbf{u}$ in the design of an attack strategy for James, and requiring that the attack yield a decoding error with constant probability on $\mathcal{C}'$, one can obtain the desired tight converse.

Acknowledgements

The work of B.K. Dey was supported in part by the Bharti Centre for Communication, IIT Bombay. The work of S. Jaggi was supported by the Research Grants Council (RGC) of Hong Kong under Project Numbers 14300617, 14304418 and 14301519. The work of M. Langberg and A.D. Sarwate was funded by the US National Science Foundation under Award CCF-1909468.

References

[1] D. Blackwell, L. Breiman, and A. J. Thomasian, "The capacities of certain channel classes under random coding," *The Annals of Mathematical Statistics*, pp. 558–567, 1960.

[2] A. Lapidoth, P. Narayan, and others, "Reliable communication under channel uncertainty," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2148–2177, 1998.

[3] I. Csiszár 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, March 1988. [Online]. Available: https://dx.doi.org/10.1109/18.2627

[4] A. D. Sarwate, "Coding against myopic adversaries," in *2010 Information Theory Workshop*, Dublin, Ireland, August-September 2010, pp. 1–5.

[5] B. K. Dey, S. Jaggi, and M. Langberg, "Sufficiently myopic adversaries are blind," in *Proceedings of the 2015 IEEE International Symposium on Information Theory (ISIT)*, Hong Kong, China, June 14–19 2015, pp. 1164–1168.

[6] ——, "Sufficiently myopic adversaries are blind," *IEEE Transactions on Information Theory*, vol. 65, no. 9, pp. 5718–5736, September 2019.

[7] M. Plotkin, "Binary codes with specified minimum distance," *IRE Transactions on Information Theory*, vol. 6, no. 4, pp. 445–450, 1960.

[8] X. Wang, A. J. Budkuley, A. Bogdanov, and S. Jaggi, "When are large codes possible for AVCs?" in *2019 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2019, pp. 632–636.

[9] B. K. Dey, S. Jaggi, M. Langberg, A. D. Sarwate, and C. Wang, "The interplay of causality and myopia in adversarial channel models," in *Proceedings of the 2019 IEEE International Symposium on Information Theory (ISIT)*, 2019, pp. 1002–1006.

[10] A. Berman and N. Shaked-Monderer, *Completely Positive Matrices*. World Scientific, 2003. [Online]. Available: https://doi.org/10.1142/5273

[11] B. K. Dey, S. Jaggi, M. Langberg, and A. D. Sarwate, "Upper bounds on the capacity of binary channels with causal adversaries," *IEEE Transactions on Information Theory*, vol. 59, no. 6, pp. 3753–3763, 2013.

[12] I. Csiszár and P. Narayan, "Capacity and decoding rules for classes of arbitrarily varying channels," *IEEE Trans. Inform. Theory*, vol. 35, pp. 752–759, 1989.

[13] I. Csiszár and J. Korner, "Broadcast channels with confidential messages," *IEEE transactions on information theory*, vol. 24, no. 3, pp. 339–348, 1978.