Shared Randomness in Arbitrarily Varying Channels

Sagnik Bhattacharya[‡] Amitalok J. Budkuley[†] Sidharth Jaggi[†]

***Abstract*—We study an adversarial communication problem where sender Alice wishes to send a message m to receiver Bob over an *arbitrarily varying channel* (AVC) controlled by a malicious adversary James. We assume that Alice and Bob share randomness K unknown to James. Using K, Alice first encodes the message m to a codeword $\mathbf{X}$ and transmits it over the AVC. James knows the message m, the (randomized) codebook and the codeword $\mathbf{X}$. James then inputs a *jamming state* $\mathbf{S}$ to disrupt communication; we assume a *state-deterministic* AVC where $\mathbf{S}$ completely specifies the channel noise. Bob receives a noisy version $\mathbf{Y}$ of codeword $\mathbf{X}$; it outputs a message estimate $\hat{m}$ using $\mathbf{Y}$ and the shared randomness K. We study AVCs, called 'adversary-weakened' AVCs here, where the availability of shared randomness strictly improves the optimum throughput or *capacity* over it than when it is not available; the *randomized coding capacity* characterizes the largest rate possible when K is unrestricted. In this work, we characterize the exact threshold for the amount of shared randomness K so as to achieve the randomized coding capacity for 'adversary-weakened' AVCs.**

We show that exactly $\log(n)$ equiprobable and independent bits of randomness, shared between Alice and Bob and unknown to adversary James, are both necessary and sufficient for achieving randomized coding capacity for 'adversary-weakened' AVCs. For sufficiency, our achievability is based on a randomized code construction which uses deterministic list codes along with a polynomial hashing technique which uses the shared randomness. Our converse, which establishes the necessity of $\log(n)$ bits of shared randomness, uses a known approach for binary AVCs, and extends it to general 'adversary-weakened' AVCs using a notion of *confusable* codewords.

Extended draft available at https://goo.gl/1pg6Bi

I. Introduction

The optimal throughput or *capacity* over a fixed point-to-point discrete memoryless channels (DMC) is well characterized (cf. [1]). It is known (cf. [2]) that the distinction between fixed or *deterministic* codes and *randomized* codes, where sender Alice and receiver Bob share randomness which allows joint randomization of the encoder-decoder pair, is irrelevant as the capacity remains the same in either case. Such shared randomness, however, often proves crucial when the channel law may not be fixed; for instance, the channel capacity under *adversarial* communication, where an adversary maliciously inputs jamming noise to disrupt communication, can be strictly larger for randomized codes over deterministic codes (cf. [2]). Furthermore, as the amount of shared randomness increases the maximum throughput possible saturates to the *randomized coding capacity*; the difference between the randomized coding capacity and deterministic coding capacity thus quantifies the maximum rate penalty when shared randomness is absent. It is well-known (at least since [3]) that $\mathcal{O}(\log(n))$ bits of shared randomness is *sufficient* to achieve randomized coding capacity. In this work, we study the exact threshold for the amount of common randomness where the aforementioned rate penalty vanishes and randomized coding capacity is achieved.

Blackwell et al. [4] first studied adversarial communication in the framework of information theory using the *arbitrarily varying channel (AVC)* model. We consider communication over state-deterministic AVCs[1], where channel noise is completely specified by the jamming noise or *state*. More formally, (cf. Fig 1) sender Alice and receiver Bob share randomness K comprising κ bits which is unknown to adversary James. Alice encodes a message $m \in [2^{nR}]$ and transmits the codeword $\mathbf{X} = \psi(m, K)$ over the state-deterministic AVC. Jammer James knows the message m and the randomized codebook; we further assume that James is *omniscient* and knows $\mathbf{X}$ non-causally. James inputs a jamming state $\mathbf{s}$ (this may be arbitrarily correlated to m, $\mathbf{X}$ and the randomized codebook) so as to disrupt the communication. Bob observes a noisy version $\mathbf{Y}$ of the transmitted codeword $\mathbf{X}$ over the state-deterministic AVC. We assume that both Alice and James have *type constraints* Γ_X and Λ_S resp.; these correspond to 'power constraints' which determine the set of feasible codeword and state vectors. We study state-deterministic AVCs where the capacity in the presence of shared randomness is strictly larger than when it is absent; these are called *'adversary-weakened'* AVCs (see Definition 1). Our interest lies in determining the exact threshold on κ so as to achieve the randomized coding capacity for the class of 'adversary-weakened' AVCs.

Related Work: Unlike [4] where shared randomness is unbounded, Ahlswede [3] showed that $\mathcal{O}(\log(n))$ (in particular, $2\log(n)$) bits of shared randomness are sufficient for achieving randomized coding capacity. His technique has been subsequently adapted for different AVC models (cf., for instance, [5], [6]); for AVCs with omniscient adversaries, however, it is not directly applicable. Langberg [7] employed a different approach using list codes and an extension of a well known result in [8] to characterize the randomized coding capacity with $\mathcal{O}(\log(n))$ bits of shared randomness for a binary AVC with an omniscient bit-flipping adversary. His approach was also used in [9], and later extended to a much wider class of AVCs in [10]. In the same work, Langberg [7] also showed that $\Omega(\log(n))$ bits are necessary for achieving the randomized coding capacity for aforementioned binary AVC. However, similar results for general AVCs do not exist; in this work we extend his result to more general

[‡]Dept. of Electrical Engineering, IIT Kanpur, India.
[†]Dept. of Information Engineering, CUHK, Hong Kong.

[1]See 'channel law' in Section II; more generally, AVCs can be *non-state-deterministic* where the noise in the channel comprises the jamming state and additional independent noise [2].

'adversary-weakened' AVCs.

Our Contribution: In this work, we investigate the threshold for the amount of shared randomness for achieving the randomized coding capacity of AVCs. We study the class of 'adversary-weakened' AVCs; these are AVCs where the deterministic coding capacity is strictly smaller than randomized coding capacity (most non-trival AVCs exhibit such behaviour and hence belong to this class). For every 'adversary-weakened' AVC, we show that $\log(n)$ equiprobable and independent bits of randomness, shared between Alice and Bob and unknown to adversary James, are necessary and sufficient to achieve the randomized coding capacity for that AVC. Unlike prior work, we show the existence of randomized codes which comprise of shared randomness with the number of shared bits *arbitrarily close* to $\log(n)$. Our converse which establishes the necessity of $\log(n)$ bits of shared randomness extends the approach in [7] (for symmetric binary AVCs) to general 'adversary-weakened' AVCs.

In Section II, we first discuss the notation and then state the problem. We present our main results in Section III. We give proofs in Sections IV and V, and make concluding remarks in Section VI.

II. Notation and Problem Setup

A. Notation

Upper case letters (e.g. X) denote random variables, lower case letters (e.g. x) represent the values taken by them, and calligraphic letters (e.g. $\mathcal{X}$) denote their alphabet. We use boldface notation to represent random vectors (e.g. $\mathbf{X}$), and the values taken by them (e.g. $\mathbf{x}$). Unless specified otherwise, the length of the vectors is n (e.g. $\mathbf{X} = (X_1, X_2, \cdots, X_n)$) and corresponds to the *block length* of operation. Let $\mathcal{P}(\mathcal{X})$ denote the set of all probability distributions defined over the set $\mathcal{X}$. Also, let $\mathcal{P}(\mathcal{X}|\mathcal{Y})$ denote the set of conditional distribution of a random variable with alphabet $\mathcal{X}$ conditioned on another random variable taking values in alphabet $\mathcal{Y}$. Let X and Y denote two random variables taking values in $\mathcal{X}$ and $\mathcal{Y}$ respectively. Then, we denote the by P_X, $P_{X,Y}$, $P_{X|Y}$ and $[P_{X,Y}]_X$, the distribution of X, the joint distribution of (X, Y), the conditional distribution of X given Y, and the marginal distribution of X under $P_{X,Y}$ respectively. Let $T_X^{(n)} \in \mathcal{P}^{(n)}(\mathcal{X})$ denote any length-n *type* of X; here $\mathcal{P}^{(n)}(\mathcal{X})$ denotes the set of all length-n types T_X defined over set $\mathcal{X}$. Unless specified otherwise, all types are under block length n. We slightly abuse notation by using the shorthand T_X when referring to an element in $\mathcal{P}^{(n)}(\mathcal{X})$. Similarly, let $T_{X,Y}$, $T_{X|Y}$ and $[T_{X,Y}]_X$ denote a joint type of (X, Y), conditional type of X given Y and marginal type of X given $T_{X,Y}$ respectively. Let $\mathcal{T}(T_X)$ denote the set of all length n sequences with type T_X. We denote type of $\mathbf{x}$ by $T_{\mathbf{x}}$, joint type of $(\mathbf{x}, \mathbf{y})$ by $T_{\mathbf{x},\mathbf{y}}$, and conditional type of $\mathbf{x}$ conditioned on $\mathbf{y}$ by $T_{\mathbf{x}|\mathbf{y}}$.

B. The Communication Setup

Consider the communication figure depicted in Fig. 1. Alice aims to reliably send a message M to a remote receiver Bob over an arbitrarily varying channel (AVC) controlled by jammer James. The AVC is specified in terms of the following: Alice's input $x \in \mathcal{X}$, James's jamming state $s \in \mathcal{S}$, output alphabet $y \in \mathcal{Y}$, Alice's input constraint $\Gamma_X \subseteq \mathcal{P}(\mathcal{X})$, James' state constraint $\Lambda_S \subseteq \mathcal{P}(\mathcal{S})$, and the state-deterministic and memoryless channel law $W_{Y|X,S}$ given in terms of the function $w(x, s)$. To simplify notation, we refer to this AVC by $\mathcal{A} = (\Gamma_X, \Lambda_S, W_{Y|X,S})$; here the alphabets $\mathcal{X}$, $\mathcal{S}$, $\mathcal{Y}$ and the block length n of operation are implicit. We assume that all alphabets are finite in size, and the sets $\Gamma_X \subseteq \mathcal{P}(\mathcal{X})$ and $\Lambda_S \subseteq \mathcal{P}(\mathcal{S})$ are convex. Furthermore, the AVC $\mathcal{A}$ is known to all parties Alice, James and Bob; we describe each of these entities now in more detail.

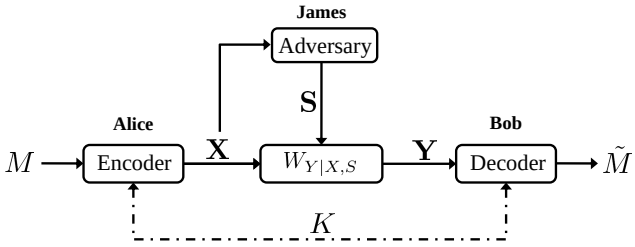

Fig. 1: The communication setup

Alice's encoder: We assume that Alice and Bob share common randomness $K \sim \text{Unif}\{[2^\kappa]\}$ (i.e., a string K comprising κ bits, chosen uniformly at random from the set $[2^\kappa]$), but unknown to James. Given the message $M = m$ and shared randomness $K = k$, Alice uses an encoder $\psi : [2^{nR}] \times [2^\kappa] \to \mathcal{X}^n$ to encode the message $M = m$ into a codeword $\mathbf{x} = \psi(m, k)$, and transmits $\mathbf{x}$ on the channel[2]; a codeword $\mathbf{x}$ is feasible if satisfies the input constraint Γ_X, i.e., $\mathbf{x}$ is such that $T_{\mathbf{x}} \in \Gamma_X$. Note that $\Gamma_X = \mathcal{P}(\mathcal{X})$ implies that Alice has an 'unconstrained input'. A *randomized* codebook $\mathcal{C} = \{\mathbf{x} = \psi(m, k) : m \in [2^{nR}], k \in [2^K]\}$ comprises all such feasible codewords under the encoder ψ; the *rate* of this code $\mathcal{C}$ is R. We note that a codebook $\mathcal{C} = \{\mathbf{x} = \psi(m, k_0) : m \in [2^{nR}]\}$, for some $k_0 \in 2^\kappa$ corresponds to a *deterministic* codebook.

James's Jammer: James chooses a (possibly random) jamming state $\mathbf{S}$ so as to disrupt communication between Alice and Bob. Apart from the AVC $\mathcal{A}$, we assume that an *omniscient* James knows the randomized code $\mathcal{C}$, the chosen message $M = m$ as well as the transmitted codeword $\mathbf{X}$; crucially, however, James does not know the randomness K shared between Alice and Bob. James' jamming state $\mathbf{S}$ can depend arbitrarily on everything that James knows. However, jamming state should satisfy its state constraint $\Lambda_S \subseteq \mathcal{P}(\mathcal{S})$, in particular, $\mathbf{S}$ should be such that $T_{\mathbf{S}} \in \Lambda_S$. James is said to have an 'unconstrained state' when $\Lambda_S = \mathcal{P}(\mathcal{S})$.

Channel law: The channel law $W^{(n)}(\mathbf{y}|\mathbf{x}, \mathbf{s})$ is specified in terms of the memoryless distribution $W_{Y|X,S}$. In particular, the probability $W^{(n)}(\mathbf{y}|\mathbf{x}, \mathbf{s}) = \prod_{i=1}^n W_{Y|X,S}(y_i|x_i, s_i)$; we

[2]We assume without loss of generality that 2^{nR} is an integer.

assume that $W_{Y|X,S}$ is *state-deterministic*, i.e., $y = w(x,s)$ for some fixed function $w : \mathcal{X} \times \mathcal{S} \to \mathcal{Y}$.

Bob's Decoder: Recall that Alice and Bob share randomness K unknown to James; let $K = k$. Upon receiving the channel output $\mathbf{y} \in \mathcal{Y}^n$, Bob uses the decoder $\phi : \mathcal{Y}^n \times [2^K] \to [2^{nR}] \cup \{0\}$ to output an estimate $\hat{m} = \phi(\mathbf{y}, k)$ of message m.

Successful communication: Given any rate R randomized code with encoder-decoder pair (ψ, ϕ), alternately called an (n, R)-*randomized code*, its corresponding probability of error error is given by

$$P_e^{(n)} := \max_m \max_{\mathbf{s}:\mathcal{X}^n \to \mathcal{S}^n} \mathbb{P}_K(\phi(\mathbf{Y}, K) \neq m | M = m)$$

where the probability is over the shared string K. A rate R is said to be *achievable* if given any $\epsilon > 0$, there exists for every n sufficiently large, an (n, R)-randomized code such that the corresponding probability of error $P_e^{(n)} \leq \epsilon$. The supremum of the all achievable rates is defined as the *randomized coding capacity* of the AVC $\mathcal{A}$; we denote it by $C_r(\mathcal{A})$. For deterministic codes, achievable rate and *deterministic coding capacity* are analogously defined; the latter is denoted by $C_d(\mathcal{A})$.

In this work, we consider the following class of 'adversary-weakened' AVCs.

Definition 1 ('Adversary-weakened' AVCs)**.** *A state-deterministic AVC $\mathcal{A} = (\Gamma_X, \Lambda_S, W_{Y|X,S}(x,s))$ is called an ''adversary-weakened' ' AVC if $C_r(\mathcal{A}) > C_d(\mathcal{A})$. We denote the class of all 'adversary-weakened' AVCs by $\mathscr{A}_{AW}$.*

Remark 1. *Most non-trivial AVCs of interest are 'adversary-weakened' AVCs where presence of shared randomness between the encoder-decoder, unknown to the adversary, allows communication at rates higher than when such shared randomness is completely absent, viz., when only deterministic codes are allowed (cf. [11] for an interesting class of AVCs where $C_d(\mathcal{A}) < C_r(\mathcal{A})$, $C_d(\mathcal{A}) = 0$). A simple example of an AVC which is not 'adversary-weakened' is as follows: let $\mathcal{X} = \mathcal{S} = \{0,1\}$, $\mathcal{Y} = \{0,1,2\}$, $\Gamma = \mathcal{P}(\mathcal{X})$, $\Gamma_X = \mathcal{P}(\mathcal{S})$, and let $y = w(x,s)$ be specified as follows: $y = x$, if $x = 0$, and $y = x + s$, if $x = 1$. It can be easily verified that this AVC is not an 'adversary-weakened' AVC.*

Our interest in this work is the threshold of the amount of shared randomness for achieving the randomized coding capacity for 'adversary-weakened' AVCs.

III. Main Results

We now state our main results. To begin, we define the following:

Definition 2. *Given an AVC $\mathcal{A} = (\Gamma_X, \Lambda_S, w(x,s))$, let*

$$\bar{C}_r(\mathcal{A}) := \max_{P_X \in \Gamma_X} \min_{P_{S|X}:[P_X P_{S|X}]_S \in \Lambda_S} I(X;Y). \quad (1)$$

Our main result is the characterization of the exact amount of shared randomness between Alice and Bob so as to achieve the randomized coding capacity for any 'adversary-weakened' AVC. In particular, we show that for any 'adversary-weakened' AVC $\mathcal{A} \in \mathscr{A}_{AW}$, exactly $\log(n)$ independent and equiprobable bits of shared randomness between Alice and Bob are both necessary and sufficient to achieve the randomized coding capacity $C_r(\mathcal{A})$ for the AVC $\mathcal{A}$. Below, we state this result.

Theorem 3. *For every 'adversary-weakened' AVC $\mathcal{A} \in \mathscr{A}_{AW}$, shared randomness K comprising $\log(n)$ bits is both necessary and sufficient for achieving its randomized coding capacity $C_r(\mathcal{A}) = \bar{C}_r(\mathcal{A})$.*

Toward proving this result, we first establish the sufficiency of this condition in Theorem 4 below.

Theorem 4 (Achievability)**.** *Consider any 'adversary-weakened' AVC $\mathcal{A} \in \mathscr{A}_{AW}$. Let $\epsilon > 0$. Then, for every n large enough, there exists an (n, R)-randomized code with shared randomness K comprising $\kappa \leq (1 + \epsilon)$ bits where rate $R = \bar{C}(\mathcal{A}) - \epsilon$.*

The proof of this theorem appears in Section IV. Next, we show that $\log(n)$ bits of shared randomness are necessary for any randomized code to achieve the randomized coding capacity for any 'adversary-weakened' AVC.

Theorem 5 (Converse)**.** *Let $\mathcal{A} \in \mathscr{A}_{AW}$. Then, every sequence of (n, R)-randomized codes with $P_e^{(n)} \to 0$ as $n \to \infty$ has rate $R \leq \bar{C}(\mathcal{A})$. Furthermore, every capacity-achieving sequence of (n, R)-randomized codes with $R = \bar{C}(\mathcal{A})$, comprises $\kappa \geq \log(n)$ bits of shared randomness K.*

We present the proof in Section V.

IV. Proof of Theorem 4

Our achievability scheme uses deterministic list codes; we begin by defining list codes formally.

Definition 6 (List code)**.** *An (n, N, L) deterministic list code is a pair of mappings (ψ_L, ϕ_L), where encoder $\psi_L : \{1, 2, \cdots, N\} \to \mathcal{X}^n$ and decoder $\phi : \mathcal{Y}^n \to N^L$. The* rate *of this code is $R = \frac{1}{n}\log(\frac{N}{L})$. For any (n, N, L) code with encoder-decoder pair (ψ_L, ϕ_L), the corresponding error[3] is given by $e^{(n)} := \max_m \max_{\mathbf{s}:\mathcal{X}^n \to \mathcal{S}^n} \mathbb{I}_{\{\Phi(\mathbf{y}) \ni m\}}$. A rate R is achievable if there exists an infinite sequence of (n, N, L) list codes of rate R with increasing block lengths n and corresponding error $e^{(n)} = 0$.*

We now show that given any AVC $\mathcal{A} \in \mathscr{A}_{AW}$, there exists a sequence of deterministic list codes with rate R arbitrarily close to $\bar{C}_r(\mathcal{A})$[4].

Lemma 7. *Given any $\epsilon > 0$, there exists a deterministic list code with rate $R = \bar{C}_r(\mathcal{A}) - \epsilon$, and list size $L = \frac{2\log(|\mathcal{Y}|)}{\epsilon}$.*

The proof uses the approach in [10]; the details can be found in [13].

We now present an outline of the proof of achievability of our main result in Theorem 4; we first describe the randomized

[3] Here $\mathbb{I}_{\{E\}}$ denotes the indicator for event E.

[4] In fact, $\bar{C}_r(\mathcal{A})$ also corresponds to the so-called *list decoding capacity* [12] of $\mathcal{A}$.

codebook construction and the concomitant encoding and decoding maps, following which we discuss the probability of error. See [13] for a detailed presentation.

Let $\epsilon > 0$. Now fix a rate $R = \bar{C}_r(\mathcal{A}) - \epsilon$. Let P_X be the optimizing distribution in (1).

Key generation: Given $\epsilon > 0$, let $\beta := \frac{3\epsilon}{4} < \epsilon$, and $l := \lceil \frac{2}{\epsilon} \rceil$. Now we define

$$\gamma := \left\lceil \frac{1+\beta}{1+l} \log(n) \right\rceil,$$

where $\gamma \in \mathbb{Z}_+$. We utilize $\gamma(l+1)$ bits; it can be verified that $\gamma(l+1) \leq (1+\epsilon)\log(n)$. Consider a field $\mathbb{F}$ of size 2^γ. Consider an integer $n \geq n_0(\epsilon) := 2^{\frac{4\epsilon+8}{\epsilon^2}}$, and generate $\gamma(l+1)$ equiprobable and independent bits. This collection of bits define the shared key K. We now split K into $(l+1)$ equal-sized chunks, i.e., $K = (K_0, K_1, \cdots, K_l)$, where each chunk $K_i \in \mathbb{F}$, $i = 0, 1, \cdots, l$. K is now revealed to Alice and Bob (though not to James).

Codebook construction: Let

$$R' := R + \frac{(1+\beta)\log(n)}{(l+1)n}. \tag{2}$$

This choice of R' is explained later in the description of the encoder. Note that $R < R' < C_r(\mathcal{A})$, where $R' \to R$ as $n \to \infty$. Lemma 7 guarantees that that there exists a deterministic list code with rate R' and list size $L = \frac{2\log(|\mathcal{Y}|)}{\epsilon}$; let $\mathcal{C}_L = (\psi_L, \phi_L)$ denote such a list code. We use $\mathcal{C}_L$ and describe the encoding and decoding in our randomized code.

Encoder: The encoder observes the message $M = m$ and key K (shared with decoder). It proceeds as follows:

• Given message $m \in [2^{nR}]$, the encoder first calculates its field-$\mathbb{F}$ representation $m^{(\mathbb{F})}$; here $m^{(\mathbb{F})} := (m_1^{(\mathbb{F})}, m_2^{(\mathbb{F})}, \cdots, m_c^{(\mathbb{F})})$, where $c := \left\lceil \frac{nR}{\gamma} \right\rceil$. and $m_i^{(\mathbb{F})} \in \mathbb{F}$, $\forall i = 1, 2, \cdots, c$.

• Next, $m^{(\mathbb{F})} \in \mathbb{F}^c$ is arranged in a l-dimensional square matrix of dimension $d := \left\lceil c^{\frac{1}{l}} \right\rceil$. Here each matrix entry is given by $m^{(\mathbb{F})}_{(i_1, i_2, \cdots, i_l)}$, where $i_j \in [d]$ for $l = 1, 2, \cdots, l$.

• The encoder now calculates a polynomial hash for the message m using shared key K given by

$$H_K(m) := K_0 + \sum_{i_1=1}^{d} \sum_{i_2=1}^{d} \cdots \sum_{i_l=1}^{d} K_1^{i_1} K_2^{i_2} \cdots K_l^{i_l} m^{(\mathbb{F})}_{(i_1, i_2, \cdots, i_l)} \tag{3}$$

using the unique collection $\{m^{(\mathbb{F})}_{(i_1, i_2, \cdots, i_l)}\}$ corresponding to message m. Here $H_K(m) \in \mathbb{F}$, $\forall m$.

• The encoder now determines the concatenated message be $\bar{M} = (m, H_K(m))$; $\bar{M} \in [2^{nR'}]$, where R' is given in (2). The encoder then transmits $\mathbf{X} = \psi_L(\bar{m})$.

Decoder: The decoder knows the shared key K and observes channel output $\mathbf{y}$. It proceeds as follows:

• Using the list decoder $\phi_L : \mathcal{Y}^n \to [2^{nR'}]^L$, the decoder outputs the following list of 'likely' candidates:

$$\mathcal{L}(\mathbf{y}) := \{i \in [2^{nR'}] : T_{\mathbf{x}_i, \mathbf{y}} \in \mathcal{T}^n \left([P_X T_{S|X} W_{Y|X,S}]_{X,Y} \right), \text{ some } T_{S|X} \text{ s.t. } [P_X T_{S|X}]_S \in \mathcal{T}^n(\mathcal{S}) \cap \Lambda_S\}.$$

• Let the collection $\{\hat{\bar{m}} = (\hat{m}, \hat{h})\}$ comprise this list $\mathcal{L}(\mathbf{y})$. If there exists a *unique* $\tilde{\bar{m}} = (\tilde{m}, \tilde{h}) \in \mathcal{L}(\mathbf{y})$, such that $(\tilde{m}, \tilde{h})$ is *consistent* w.r.t. the observed shared key K, i.e., $\tilde{h} = H_K(\tilde{m})$, then the decoder outputs the corresponding message $\tilde{m} \in [2^{nR}]$ as the estimate. Otherwise, it outputs $\tilde{m} = 0$ to declare error.

Probability of error analysis: We present an overview of the probability of error analysis. Note that decoding error occurs if at least one of the following events occur:

(a) list $\mathcal{L}(\mathbf{y})$ output by the decoder does not contain $\bar{m} = (m, H_K(m))$ which corresponds to the actual message m.

(b) there exists $\bar{m}' \in \mathcal{L}(\mathbf{y})$, $\bar{m}' = (m', H_K(m'))$, where $m' \neq m$ and $H_K(m')$ is *consistent* with the observed key K (cf. (3)).

• We analyse a *stronger* adversary James by revealing to it the hash $H_K(m)$ corresponding to the message m; recall that James already knows m, the codebook and the transmitted codeword $\mathbf{X}$.

• Recall $\mathcal{C} = (\psi_L, \phi_L)$ from earlier; it immediately follows that the probability of the error event in part (a) is zero.

• The analysis of the second part is considerably more involved. To begin, we further *strengthen* the adversary by allowing it to induce 'arbitrary' lists $\mathcal{L}(\mathbf{y})$ at the decoder as long as the actual message $\bar{m} = (m, H_K(m)) \in \mathcal{L}(\mathbf{y})$ and the list size $|\mathcal{L}(\mathbf{y})| \leq \frac{2\log(|\mathcal{Y}|)}{\epsilon}$. Henceforth, we analyse such a modified list $\mathcal{L}(\mathbf{y})$.

• Thus, error occurs if some $\bar{m}' = (m', H_K(m')) \in \mathcal{L}(\mathbf{y})$, $m \neq m$, with the corresponding hash (consistent with K) $H_K(m') = H_K(m)$ (recall that $\bar{m} \in \mathcal{L}(\mathbf{y})$). To bound the corresponding error probability, we leverage the fact that even under James' complete knowledge (which includes the hash $H_K(m)$), there is still 'enough' uncertainty about the shared key K. Recall that $K = (K_0, K_1, \cdots, K_l)$; we show in [13,] that $(K_1, K_2, \cdots, K_l)$, conditioned the James' entire knowledge, is uniformly distributed. We then bound the probability of the event $H_K(m') = H_K(m)$, $m' \neq m$, conditioned on James' knowledge. Note that this corresponds determining the probability that some polynomial in $(K_1, K_2, \cdots, K_l)$ evaluates to zero (cf. (3); we use the Schwartz-Zippel lemma [14], [15] to bound this probability.

• As $|\mathcal{L}(\mathbf{y})| \leq \frac{2\log(|\mathcal{Y}|)}{\epsilon}$, $\forall \mathbf{y}$, there can exist at most $\frac{2\log(|\mathcal{Y}|)}{\epsilon}$ such candidates $\bar{m}'$. We taking a union bound to obtain an upper bound on the probability of error for part (b).

• This gives us an overall probability of error $P_e^{(n)} \leq c_1 n^{-c_2}$, where $c_1, c_2 > 0$ and independent of n, which is vanishing as $n \to \infty$. As $\epsilon > 0$ is arbitrary, this completes the proof of achievability.

V. Proof of Converse for Theorem 4

We prove a converse result under the *average* probability of error criterion instead of the *maximum* probability of error criterion. Given an (n, R)-randomized code, we define the corresponding average probability of error by $P_e^{(n)} := \frac{1}{2^{nR}} \sum_{m=1}^{2^{nR}} \max_{\mathbf{s}: \mathcal{X}^n \to \mathcal{S}^n} \mathbb{P}_K(\phi(\mathbf{Y}, K) \neq m | M = m)$. For the rest of this section, we assume $P_e^{(n)}$ to denote this average

probability of error. Furthermore, we assume a 'weaker' adversary, where (unlike in the problem description) the adversary knows only the transmitted codeword and the randomized code, but *not* the message chosen; these assumptions result in a 'stronger'[5] converse.

Consider any sequence (with increasing block length), say $\{\mathcal{C}_n\}$, of (n, R)-randomized codes. We prove the following:

1) If $\{\mathcal{C}_n\}$ is such that the corresponding $P_e^{(n)} \to 0$ as $n \to \infty$, then, its rate $R \leq C_r(\mathcal{A})$; observe that this bound on the rate holds irrespective of the amount of shared randomness. This proof is along the lines of the converse of the standard discrete memoryless channel [16] with a few modification; we defer the details to [13].
2) Let the sequence of codes $\{\mathcal{C}_n\}$ be capacity-achieving, i.e., let $R = C_r$. Then, the shared randomness K comprises at least $\log(n)$ independent and equiprobable bits, i.e., $\kappa \geq \log(n)$.

Our proof for the second part is based upon the approach in [7]; we show the necessity of $\log(n)$ bits of shared randomness through a proof via a contradiction. In particular, given $\mathcal{A} \in \mathscr{A}_{AW}$, let us consider any sequence of (n, R)-randomized codes $\{\mathcal{C}_n\}$, each comprising $\kappa < \log(n)$ bits of shared randomness, with $R = C_r(\mathcal{A})$ and average probability of error $P_e^{(n)} < \delta$. In our proof, we construct a feasible (w.r.t. the state constraint set Λ_S) jamming strategy which results in the corresponding average probability of error being at least δ, i.e., $P_e^{(n)} \geq \delta$, thereby making the contradiction (recall that we assumed $P_e^{(n)} < \delta$) and establishing the necessity of $\log(n)$ bits of shared randomness.

We now present an overview of the proof and emphasize the key ideas; the detailed proof is given in [13].

• We first determine the set $\mathcal{U} \subseteq \mathcal{X}^n$ comprising *all* codewords for the given randomized code (cf. [13, Definition 13]).

• Within this set, we seek pairs of *confusable* codewords (see [13, Definition 17]). Roughly speaking, two codewords $\mathbf{x}$, $\mathbf{x}'$ are said to be *confusable* if there exist corresponding feasible jamming states $\mathbf{s}(\mathbf{x}), \mathbf{s}'(\mathbf{x}')$ (i.e., $T_{\mathbf{s}}, T_{\mathbf{s}'} \in \Lambda_S$) which result in an identical channel output $\mathbf{y}$, i.e., $\mathbf{y} = w(\mathbf{x}, \mathbf{s}) = w(\mathbf{x}', \mathbf{s}')$; observe that *every* decoder is *confused* under such a $\mathbf{y}$ (both $\mathbf{x}, \mathbf{x}'$ are equally likely to have caused $\mathbf{y}$).

• A key part of the proof involves showing that there exist 'many' pairs of 'confusable codewords' (cf. [13, Lemma 20]); furthermore, such confusable pairs occur in 'many' subcodebooks (each instantiated by K) in the randomized code ensemble (cf. [13, Lemma 16]). The proof of [13, Lemma 20] crucially uses the fact that the underlying channel corresponds to an 'adversary-weakened' AVC, where $C_d(\mathcal{A}) < C_r(\mathcal{A})$.

• Given the randomized codebook, say $\mathcal{C}$, we now propose the following jamming attack for James: for every pair of confusable codewords $\mathbf{x}, \mathbf{x}' \in \mathcal{C}$, James chooses corresponding state vectors $\mathbf{s}, \mathbf{s}'$ such that $w(\mathbf{x}, \mathbf{s}) = w(\mathbf{x}', \mathbf{s}') = \mathbf{y}$, for some $\mathbf{y} \in \mathcal{Y}^n$; note that we are guaranteed that there exist such feasible $\mathbf{s}, \mathbf{s}'$ and output $\mathbf{y}$. For codewords $\mathbf{x} \in \mathcal{C}$, which do not comprise any confusable pair, James chooses $\mathbf{s} = 0$. This determines James' strategy via the collection $\{\mathbf{s}(\mathbf{x}); \mathbf{x} \in \mathcal{C}\}$.

• As every confusable pair causes a decoding error and since a 'large' fraction of such pairs of confusable codewords occur, a calculation (cf. [13]) then shows that the average probability of error is at least δ, i.e., $P_e^{(n)} \geq \delta$. This establishes the contradiction.

[5]This is because derived bounds on rate and the size of the common randomness under the proposed 'weakening' of the adversary as well as the error criterion continue to hold even when the model is restored back to one in the problem description.

VI. Conclusion

We study the communication problem in the presence of an adversary when the encoder-decoder can share randomness which is not revealed to the adversary. Randomized coding capacity is the optimum throughput when shared randomness of arbitrary size is available. We characterize the exact threshold on the amount of shared randomness required for realizing the randomized coding capacity for a large class of 'adversary-weakened' AVCs. In particular, we show that $\log(n)$ bits of shared randomness are necessary as well as sufficient to achieve randomized coding capacity for any 'adversary-weakened' AVC.

References

[1] C. Shannon, "A mathematical theory of communication," *Bell Sys. Tech. Journal*, vol. 27, 1948.

[2] A. Lapidoth and P. Narayan, "Reliable communication under channel uncertainty," *IEEE Trans. Inform. Theory*, vol. 44, pp. 2148–2177, October 1998.

[3] R. Ahlswede, "Elimination of correlation in random codes for arbitrarily varying channels," *Z. Wahrscheinlichkeitstheorie Verv. Gebiete*, vol. 44, pp. 181–193, 1978.

[4] D. Blackwell, L. Breiman, and A. J. Thomasian, "The capacity of a class of channels," *Ann. of Mathematical Statistics*, vol. 30, no. 4, pp. 1229–1241, 1959.

[5] R. Ahlswede, "Arbitrarily varying channels with states sequence known to the sender," *IEEE Trans. Inform. Theory*, vol. 32, pp. 621–629, September 1986.

[6] S. Watanabe and S. Kuzuoka, "Universal Wyner-Ziv coding for distortion constrained general side information," *IEEE Trans. Inform. Theory*, vol. 60, pp. 7568–7583, December 2014.

[7] M. Langberg, "Private codes or succinct random codes that are (almost) perfect," in *Proc. IEEE Int. Symp. Found. Comp. Sci*, Rome, Italy, 2004.

[8] P. Erdös, P. Frankl, and Z. Füredi, "Families of finite sets in which no set is covered by the union ofr others," *Israel Journal of Mathematics*, vol. 51, no. 1, pp. 79–89, 1985.

[9] A. Smith, "Scrambling adversarial errors using few random bits, optimal information reconciliation, and better private codes," in *Proc. ACM-SIAM Symp. on Discrete Algorithms*, vol. 7, no. 09. Citeseer, 2007, pp. 395–404.

[10] A. Sarwate, "Robust and adaptive communication under uncertain interference," Ph.D. dissertation, University of California, Berkeley, 2008.

[11] X. Wang, A. J. Budkuley, A. Bogdanov, and S. Jaggi, "When are large codes possible for AVCs?" *Preprint*, 2019. [Online]. Available: http://goo.gl/2saETh

[12] R. Ahlswede, "Channel capacities for list codes," *Journ. Appl. Prob.*, vol. 10, pp. 824–836, 1973.

[13] S. Bhattacharya, A. J. Budkuley, and S. Jaggi, "Shared randomness in arbitrarily varying channels," *Preprint*, 2019. [Online]. Available: https://goo.gl/1pg6Bi

[14] J. T. Schwartz, "Fast probabilistic algorithms for verification of polynomial identities," *Journal of ACM*, vol. 27, pp. 701–717, 1980.

[15] R. Zippel, "Probabilistic algorithms for sparse polynomial," in *Proc. Int. Symp. Symb. Alg. Comp.*, Hong Kong, China, June 1979.

[16] A. E. Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge University Press, 2011.