The Capacity of Causal Adversarial Channels

Yihan Zhang
Institute of Science and Technology Austria
zephyr.z798@gmail.com

Sidharth Jaggi
University of Bristol
sid.jaggi@bristol.ac.uk

Michael Langberg
University at Buffalo
mikel@buffalo.edu

Anand D. Sarwate
Rutgers University
anand.sarwate@rutgers.edu

2022 IEEE International Symposium on Information Theory (ISIT) | 978-1-6654-2159-1/22/$31.00 ©2022 IEEE | DOI: 10.1109/ISIT50566.2022.9834709

***Abstract*—We characterize the capacity for the discrete-time arbitrarily varying channel with discrete inputs, outputs, and states when (a) the encoder and decoder do not share common randomness, (b) the input and state are subject to cost constraints, (c) the transition matrix of the channel is deterministic given the state, and (d) at each time step the adversary can only observe the current and past channel inputs when choosing the state at that time. The achievable strategy involves stochastic encoding together with list decoding and a disambiguation step. The converse uses a two-phase "babble-and-push" strategy where the adversary chooses the state randomly in the first phase, list decodes the output, and then chooses state inputs to symmetrize the channel in the second phase. These results generalize prior work on specific channels models (additive, erasure) to general discrete alphabets and models.**

***Index Terms*—arbitrarily varying channels, channel capacity, jamming**

I. Introduction

In introductory courses on information theory and coding theory students encounter two basic models for communication channels. The Shannon-theoretic model [1] for memoryless channels treats the effect of the channel as random, where each input symbol is transformed to an output symbol through the same conditional distribution at each time step. Two canonical examples are the binary symmetric channel (BSC) and binary erasure channel (BEC). With high probability, for sufficiently large n, a BSC flips close to pn bits for a codeword of blocklength n and the probability of error is *average-case*, measured over the randomness in the channel. By contrast, in the basic coding theory model, errors and erasures are modeled as *worst-case*: for a blocklength n the goal is to design a code which can correct any pattern of pn errors or erasures.

One way to understand the difference between these models is to frame them both in the context of arbitrarily varying channels (AVCs) [2] under constraints [3], [4]. In the AVC there are three participants: Alice (the transmitter/encoder), Bob (the receiver/decoder), and James (an adversarial jammer). When communicating over an AVC, Alice encodes her message into a codeword $\underline{x}$ of blocklength n and James can choose an equal-length vector of channel states $\underline{s}$. The output $\underline{y}$ is formed by applying a channel law $W_{\mathbf{y}|\mathbf{x},\mathbf{s}}(y|x,s)$ letter-by-letter to $(\underline{x},\underline{s})$. The difference between the two classical communication models can be captured by modeling the information James has about the transmitted codeword. The Shannon-theoretic model is similar to an *oblivious adversary* who must choose $\underline{s}$ without any knowledge of $\underline{x}$. The coding-theoretic model is similar to a *omniscient adversary* in which James can choose $\underline{s}$ as a function of the entire codeword $\underline{x}$.

Once we frame the difference between average and worst case models in terms of the AVC, a variety of "intermediate case" models become natural by changing what James can know about the transmitted codeword. In this paper we consider one such model: the *causal (or online) adversary* in which James chooses the channel state $\underline{s}(t)$ at time t based on knowledge of the current and past inputs $(\underline{x}(1),\underline{x}(2),\ldots,\underline{x}(t))$. The online adversary is a special case of the *delayed adversary* [5], [6] who generates each state symbol $\underline{s}(t)$ based on the delayed observations $(\underline{x}(1),\underline{x}(2),\cdots,\underline{x}(t-\Delta))$ for some integer $0 \leqslant \Delta \leqslant t$.

Much of the prior work on causal adversaries deals with specific channel models. Capacity results for special cases of causal adversaries with "large alphabets" [7], the erasure setting [8], [9], [10], the bit-flip/symbol-error setting [11], [10], and the quadratically-constrained scenario [12] are known. Other related channel models explored include settings with a memoryless jammer [13], and bit-flip and erasure models in which the channel is not state-deterministic but James can observe the channel output [14].

In this work we focus on AVCs with finite input, state, and output alphabets which are *state-deterministic*, meaning the channel output $\underline{y}(t)$ at each time t is a deterministic function of $\underline{x}(t)$ and $\underline{s}(t)$. In such models, James can compute the channel output. Our goal is to establish capacity results for general state-deterministic AVC models with cost constraints.

After defining our model in Section II, and key concepts in Section III, we present an overview of our capacity analysis in Section IV. At a high level, both our achievability and converse proofs follow those appearing in [11], [10] addressing the causal bit-flip/symbol-error setting. The main technical contribution in this work thus lies in the nature of expanding the concepts and analysis in [11], [10] to fit the generalized model of AVCs with both state and input constraints (cf. Section IV). We highlight the major challenges of analyzing general AVCs and the tools used to overcome these challenges in the overview of Section IV. Formal proofs are provided in an extended version of this manuscript [15].

II. Model

Notation: In this paper, all alphabets are finite and in calligraphic script (e.g. $\mathcal{X}$). A boldface letter (e.g. $\mathbf{x}$) in-

The work of ADS and ML was supported in part by the US National Science Foundation under awards CCF-1909468 and CCF-1909451.

dicates a random variable, and non-boldface (e.g. x) as its realization. The set $[M] = \{1, 2, \ldots, M\}$. Tuples are written with an underline sign and individual entries with the time index in parentheses (e.g. $\underline{x} = (\underline{x}(1), \underline{x}(2), \ldots, \underline{x}(n))$) and $\underline{x}(1:i) = (\underline{x}(1), \underline{x}(2), \ldots, \underline{x}(i))$). The *type* $T_{\underline{x}}$ of a tuple $\underline{x}$ is the empirical distribution of $\underline{x}$. The set of all probability distributions on an alphabet $\mathcal{X}$ is $\Delta(\mathcal{X})$. The set of all conditional distributions (randomized maps) from $\mathcal{X}$ to $\mathcal{S}$ is $\Delta(\mathcal{S}|\mathcal{X})$. The length n *type class* corresponding to $P \in \Delta(\mathcal{X})$ is denoted by $\mathcal{T}_n(P) = \left\{\underline{x} \in \mathcal{X}^n : T_{\underline{x}} = P\right\}$. For a joint distribution $P_{\mathbf{x},\mathbf{s}}$ we write $[P_{\mathbf{x},\mathbf{s}}]_{\mathbf{x}}$ and $[P_{\mathbf{x},\mathbf{s}}]_{\mathbf{s}}$ for the marginal distributions of $\mathbf{x}$ and $\mathbf{s}$.

A. Channels and codes

We consider a class of arbitrarily varying channels (AVCs) with cost constraints on the input and state. Our formulation of the cost constraint generalizes the standard definition [3] by modeling the constraint as requiring that the type of the channel input or state belong to a specified set.

Definition 1 (AVC)**.** An *arbitrarily varying channel* (AVC) is a sextuple $(\mathcal{X}, \mathcal{S}, \mathcal{Y}, \lambda_{\mathbf{x}}, \lambda_{\mathbf{s}}, W_{\mathbf{y}|\mathbf{x},\mathbf{s}})$. Here, $\mathcal{X}, \mathcal{S}, \mathcal{Y}$ are the input, state and output alphabets, respectively. The input and state constraints are specified by $\lambda_{\mathbf{x}} \subset \Delta(\mathcal{X})$ and $\lambda_{\mathbf{s}} \subset \Delta(\mathcal{S})$, respectively. The channel law is $W_{\mathbf{y}|\mathbf{x},\mathbf{s}} \in \Delta(\mathcal{Y}|\mathcal{X} \times \mathcal{S})$.

Our goal is to communicate one of M messages reliably over this AVC. For a positive integer M, let $\mathcal{M} := [M]$ denote all possible messages that the transmitter may send.

Definition 2 (Codes)**.** A *code* for a causal AVC $(\mathcal{X}, \mathcal{S}, \mathcal{Y}, \lambda_{\mathbf{x}}, \lambda_{\mathbf{s}}, W_{\mathbf{y}|\mathbf{x},\mathbf{s}})$ is a pair $(\mathrm{Enc}, \mathrm{Dec})$. Here $\mathrm{Enc} \in \Delta(\mathcal{X}^n|\mathcal{M})$ is a (potentially stochastic) encoder. For $m \in \mathcal{M}$, we use $\mathrm{Enc}(m) \in \mathcal{X}^n$ to denote the (possibly random) encoding of m. Each such an encoding is called a *codeword*. The set $\mathrm{Enc}(\mathcal{M})$ of all codewords is called the *codebook*, denoted by $\mathcal{C}$. The length n of each codeword is called the *blocklength*. The *rate* of $\mathcal{C}$ is defined as $R(\mathcal{C}) := \frac{1}{n} \log M$. The code is required to satisfy the input constraint: for every $\underline{x} \in \mathcal{C}$, $T_{\underline{x}} \in \lambda_{\mathbf{x}}$. The decoder is given by $\mathrm{Dec} \in \Delta(\mathcal{M}|\mathcal{Y}^n)$. We use $\mathrm{Dec}(\underline{y}) \in \mathcal{M}$ to denote the (potentially random) message output by the decoder given $\underline{y}$.

Definition 3 (Jamming strategies)**.** A *jamming strategy* of blocklength n is a set of maps $\mathrm{Jam} = (\mathrm{Jam}_1, \cdots, \mathrm{Jam}_n)$ where $\mathrm{Jam}_t \in \Delta(\mathcal{S}|\mathcal{X}^n)$ is the jamming function at time t. In a *causal jamming strategy*, $\mathrm{Jam}_t \in \Delta(\mathcal{S}|\mathcal{X}^t)$.

Definition 4 (Communication over causal/online AVC)**.** Communication over a *causal* (a.k.a. *online*) AVC $(\mathcal{X}, \mathcal{S}, \mathcal{Y}, \lambda_{\mathbf{x}}, \lambda_{\mathbf{s}}, W_{\mathbf{y}|\mathbf{x},\mathbf{s}})$ has the following requirements. Let $\mathcal{C}$ be a code with an encoder-decoder pair $(\mathrm{Enc}, \mathrm{Dec})$ and $\mathrm{Jam} = (\mathrm{Jam}_1, \cdots, \mathrm{Jam}_n)$ be a causal jamming strategy. We use $\mathrm{Jam}_i(\underline{x}(1), \cdots, \underline{x}(i))$ to denote the jamming symbol generated by the adversary at time i. Before communication happens, $(\mathrm{Enc}, \mathrm{Dec})$ are fixed and revealed to the transmitter *Alice*, the receiver *Bob* and the adversary *James*. James then fixes a causal jamming strategy $\mathrm{Jam} = (\mathrm{Jam}_1, \cdots, \mathrm{Jam}_n)$ which can depend on $(\mathrm{Enc}, \mathrm{Dec})$.

The code is required to satisfy the input constraint $T_{\underline{x}} \in \lambda_{\mathbf{x}}$ for every $\underline{x} \in \mathcal{C}$. Once a particular encoding $\underline{x}$ of a certain message m is transmitted by Alice, James observes $\underline{x}$ *causally*. That is, for any $i \in [n]$, he observes $\underline{x}(i)$ after observing $\underline{x}(1), \cdots, \underline{x}(i-1)$. Given his causal observation, he computes $\underline{s}(i) = \mathrm{Jam}_i(\underline{x}(1), \cdots, \underline{x}(i))$ which depends only on $\underline{x}(1), \cdots, \underline{x}(i)$ (and $(\mathrm{Enc}, \mathrm{Dec}), \mathcal{C}$ which are known to everyone). The channel then outputs $\underline{\mathbf{y}}$ according to the following distribution

$$\Pr\left[\underline{\mathbf{y}} = \underline{y} \,\middle|\, \underline{\mathbf{x}} = \underline{x}, \underline{\mathbf{s}} = \underline{s}\right] := \prod_{i=1}^{n} W_{\mathbf{y}|\mathbf{x},\mathbf{s}}(\underline{y}(i)|\underline{x}(i), \underline{s}(i)).$$

Receiving $\underline{\mathbf{y}}$, Bob decodes to $\mathrm{Dec}(\underline{\mathbf{y}})$.

We consider the maximum (resp. average) probability of error criterion in achievability (resp. converse).

Definition 5 (Error probability)**.** Consider a causal AVC $(\mathcal{X}, \mathcal{S}, \mathcal{Y}, \lambda_{\mathbf{s}}, \lambda_{\mathbf{s}}, W_{\mathbf{y}|\mathbf{x},\mathbf{s}})$. Let $(\mathrm{Enc}, \mathrm{Dec})$ and Jam be a coding scheme and a jamming strategy for this channel, respectively. Define the *maximum error probability* $P_{\mathrm{e,max}}(\mathrm{Enc}, \mathrm{Dec}, \mathrm{Jam})$ as follows (the *average error probability* $P_{\mathrm{e,avg}}$ is defined analogously by averaging over messages m).

$$\max_{m \in \mathcal{M}} \sum_{\substack{\widehat{m} \in \mathcal{M} \\ \widehat{m} \neq m}} \sum_{\underline{y} \in \mathcal{Y}^n} \sum_{\underline{s} \in \mathcal{S}^n} \sum_{\underline{x} \in \mathcal{X}^n} \mathrm{Enc}(\underline{x}|m) \cdot$$

$$\left(\prod_{i=1}^{n} \mathrm{Jam}_i(\underline{s}(i)|\underline{x}(1:i)) W_{\mathbf{y}|\mathbf{x},\mathbf{s}}(\underline{y}(i)|\underline{x}(i), \underline{s}(i)) \right) \cdot \mathrm{Dec}(\widehat{m}|\underline{y}).$$

Here we view $\mathrm{Enc} \in \Delta(\mathcal{X}^n|\mathcal{M})$, $\mathrm{Jam}_i \in \Delta(\mathcal{S}|\mathcal{X}^i)$ and $\mathrm{Dec} \in \Delta(\mathcal{M}|\mathcal{Y}^n)$ as conditional distributions.

B. State-deterministic AVCs

We consider a class of AVCs which are state deterministic and have a single cost constraint on the state. More precisely, we require the following five assumptions to hold:

1) All alphabets $\mathcal{X}, \mathcal{S}, \mathcal{Y}$ are finite.[1]
2) The input constraint set $\lambda_{\mathbf{x}} \subset \Delta(\mathcal{X})$ is convex. This is a natural restriction – a non-convex set $\lambda_{\mathbf{x}}$ would imply that the encoder is not allowed to time-share between some potential transmissions.
3) The set $\lambda_{\mathbf{s}} \subset \Delta(\mathcal{S})$ is specified by a *single* constraint:
$$\lambda_{\mathbf{s}} := \left\{ P_{\mathbf{s}} \in \Delta(\mathcal{S}) : \sum_{s \in \mathcal{S}} P_{\mathbf{s}}(s) B(s) \leqslant \Lambda \right\},$$
for some $B \in \mathbb{R}^{|\mathcal{S}|}$ and $\Lambda \in \mathbb{R}$. This assumption will be used in the achievability proof (cf. [15, Claim 9]).
4) The channel law $W_{\mathbf{y}|\mathbf{x},\mathbf{s}}$ is deterministic, i.e., for every $(x, s) \in \mathcal{X} \times \mathcal{S}$, there is a unique $y \in \mathcal{Y}$ such that $W_{\mathbf{y}|\mathbf{x},\mathbf{s}}(y|x, s) = 1$. Alternatively, we write the channel law as a (deterministic) function $W : \mathcal{X} \times \mathcal{S} \to \mathcal{Y}$. [2]

[1] The quadratically-constrained infinite alphabet setting was considered in prior work [12].

[2] Our achievability result actually does not require this restriction – it holds even if the AVC is not state-deterministic. However, our current converse arguments providing a capacity upper bound asymptotically matching the rate achievable by our achievability scheme rely on state-determinism, since they rely on the jammer being able to predict the channel output resulting from a specific jamming strategy.

5) There exists a zero-cost state $s_0 \in \mathcal{S}$ and a one-to-one mapping $\phi : \mathcal{X} \to \mathcal{Y}$ for which for every x, $W_{\mathbf{y}|\mathbf{x},\mathbf{s}}(\phi(x)|x, s_0) = 1$. This final assumption is rather natural and corresponds to many channel models. It intuitively implies that there is a *stand-down* state for James that on one hand has zero cost and on the other does not corrupt communication at all, e.g., the "non-erasure" state in an erasure channel. This assumption will be used in [15, Claim 9].

III. Main Result and Symmetrization

In order to state our main result we must define a notion of *symmetrizability* [16], [4] that is appropriate to the causal AVC model. Symmetrizabillity conditions play an important role in characterizing AVC capacities under deterministic coding. Roughly speaking, a channel is symmetrizable if James can, via selecting the state sequence $\underline{s}$, cause the channel to behave like a symmetric two-user multiple-access channel $W_{\mathbf{y}|\mathbf{x},\mathbf{x}'}(y|x,x')$. Operationally, this means James can select an alternative ("spoofing") codeword $\underline{x}'$ and use it to create a state $\underline{s}$ such that Bob cannot tell if Alice sent $\underline{x}$ and James chose $\underline{x}'$ or if Alice sent $\underline{x}'$ and James chose $\underline{x}$.

There are two aspects of the causal AVC which make it tricky to define a notion of symmetrizability. First is the online nature of the adversarial attack: James has to choose the elements of $\underline{s}$ sequentially as opposed to selecting $\underline{x}'$ and then $\underline{s}$. Second is the cost constraint on $\underline{s}$: if we think of the cost as "power," then James faces a power allocation problem. Taken together, James could spend little power at the beginning and more power at the end of the transmission or vice versa. In the former case, Bob can get a good estimate of the message initially but then the channel becomes much worse. In the latter, Bob has a very bad estimate of the message but the channel is less noisy at the end, allowing him to potentially decode to the true message. Since James does not know the transmitted codeword a priori, he has to choose how to allocate the power "on the fly" while satisfying the cost constraint.

A. Symmetrizing distributions

Let K be a positive integer, $\mathcal{U} := [K]$ and $\mathcal{A} := \{0, \frac{1}{K}, \frac{2}{K}, \dots, 1 - \frac{1}{K}\}$. For $\alpha \in \mathcal{A}$, let $\mathcal{U}^{\leqslant\alpha} := [\alpha K]$ and $\mathcal{U}^{>\alpha} := [K] \backslash [\alpha K]$. Let $P_{\mathbf{u}} \in \Delta(\mathcal{U})$ be the uniform distribution. We define $P_{\mathbf{x}|\mathbf{u}} \in \Delta(\mathcal{X}|\mathcal{U})$ such that $\left[P_{\mathbf{u}} P_{\mathbf{x}|\mathbf{u}}\right]_{\mathbf{x}} \in \lambda_{\mathbf{x}}$.

In our achievable scheme we will use a code in which the total blocklength n is broken into K subblocks (which we call *chunks*). We consider encoding and decoding strategies that operate in two phases, the first having αK chunks and other having $(1-\alpha)K$ chunks. In each chunk, we characterize James's actions by a single letter channel. In the first phase, James chooses $\mathbf{s}$ based on the the input $\mathbf{x}$ and chunk $\mathbf{u}$ using a channel $V_{\mathbf{s}|\mathbf{x},\mathbf{u}^{\leqslant\alpha}} \in \Delta(\mathcal{S}|\mathcal{X} \times \mathcal{U}^{\leqslant\alpha})$, and in the second phase James chooses $\mathbf{s}$ based on the input $\mathbf{x}$, $\mathbf{u}$, and an alternative *spoofing* input $\mathbf{x}'$ using the channel $V_{\mathbf{s}|\mathbf{x},\mathbf{x}',\mathbf{u}^{>\alpha}} \in \Delta(\mathcal{S}|\mathcal{X}^2 \times \mathcal{U}^{>\alpha})$. Our distinction between the two modes of James' operation supports our converse "babble-and-push" proof paradigm (see Section IV-A) in which James first generates his jamming state s through $V_{\mathbf{s}|\mathbf{x},\mathbf{u}^{\leqslant\alpha}}$ and then tries to symmetrize using a spoofing codeword $\underline{x}'$ through $V_{\mathbf{s}|\mathbf{x},\mathbf{x}',\mathbf{u}^{>\alpha}}$. We note that this distinction does not limit James in any way once we address achievability.

We define the induced single letter distribution over $\mathcal{S}$ corresponding to such a strategy by

$$Q(V_{\mathbf{s}|\mathbf{x},\mathbf{u}^{\leqslant\alpha}}, V_{\mathbf{s}|\mathbf{x},\mathbf{x}',\mathbf{u}^{>\alpha}}) := \frac{1}{\alpha K} \sum_{u=1}^{\alpha K} \left[P_{\mathbf{x}|\mathbf{u}=u} V_{\mathbf{s}|\mathbf{x},\mathbf{u}^{\leqslant\alpha}=u}\right]_{\mathbf{s}} + \frac{1}{(1-\alpha)K} \sum_{u=\alpha K+1}^{K} \left[P_{\mathbf{x}|\mathbf{u}=u}^{\otimes 2} V_{\mathbf{s}|\mathbf{x},\mathbf{x}',\mathbf{u}^{>\alpha}=u}\right]_{\mathbf{s}}.$$

Note that importantly the second term is computed according to the product distribution $P_{\mathbf{x}|\mathbf{u}=u}^{\otimes 2}$. If $V_{\mathbf{s}|\mathbf{x},\mathbf{x}',\mathbf{u}^{>\alpha}}$ does not depend on $\mathbf{x}'$ then we can define $Q(V_{\mathbf{s}|\mathbf{x},\mathbf{u}}) := \frac{1}{K}\sum_{u=1}^{K}\left[P_{\mathbf{x}|\mathbf{u}=u} V_{\mathbf{s}|\mathbf{x},\mathbf{u}=u}\right]_{\mathbf{s}}$ as a special case. We can now define two sets of feasible jamming strategies, i.e., two sets of distributions that satisfy the cost constraint.

Definition 6 (Feasible jamming distributions)**.** Let $P_{\mathbf{x}|\mathbf{u}} \in \Delta(\mathcal{X}|\mathcal{U})$. Define, for $\alpha \in \mathcal{A}$,

$$\mathcal{F}_\alpha(P_{\mathbf{x}|\mathbf{u}}) = \{(V_{\mathbf{s}|\mathbf{x},\mathbf{u}^{\leqslant\alpha}}, V_{\mathbf{s}|\mathbf{x},\mathbf{x}',\mathbf{u}^{>\alpha}}) : Q(\cdot) \in \lambda_{\mathbf{s}}\}, \quad (1)$$

$$\mathcal{F}(P_{\mathbf{x}|\mathbf{u}}) = \{V_{\mathbf{s}|\mathbf{x},\mathbf{u}} : Q(\cdot) \in \lambda_{\mathbf{s}}\}. \quad (2)$$

Definition 7 (Cumulative mutual information)**.** Fix $P_{\mathbf{x}|\mathbf{u}}$, $\alpha \in \mathcal{A}$ and $V_{\mathbf{s}|\mathbf{x},\mathbf{u}^{\leqslant\alpha}} \in \Delta(\mathcal{S}|\mathcal{X} \times \mathcal{U}^{\leqslant\alpha})$. The *cumulative mutual information* w.r.t. $P_{\mathbf{x}|\mathbf{u}}$ and $V_{\mathbf{s}|\mathbf{x},\mathbf{u}^{\leqslant\alpha}}$ is defined as $I(P_{\mathbf{x}|\mathbf{u}}, V_{\mathbf{s}|\mathbf{x},\mathbf{u}^{\leqslant\alpha}}) := I(\mathbf{x};\mathbf{y}|\mathbf{u}^{\leqslant\alpha}) = \frac{1}{K}\sum_{u=1}^{\alpha K} I(\mathbf{x}_u;\mathbf{y}_u)$, where the joint distribution of $(\mathbf{x}_u, \mathbf{y}_u)$ is given by

$$P_{\mathbf{x}_u,\mathbf{y}_u}(x,y) := \sum_{s\in\mathcal{S}} P_{\mathbf{x}|\mathbf{u}}(x|u) V_{\mathbf{s}|\mathbf{x},\mathbf{u}^{\leqslant\alpha}}(s|x,u) W_{\mathbf{y}|\mathbf{x},\mathbf{s}}(y|x,s).$$

$I(P_{\mathbf{x}|\mathbf{u}}, V_{\mathbf{s}|\mathbf{x},\mathbf{u}^{\leqslant\alpha}})$ represents the normalized amount of information reaching Bob in the first αK chunks of the transmitted codeword $\underline{x}$ under James attack governed by $V_{\mathbf{s}|\mathbf{x},\mathbf{u}^{\leqslant\alpha}}$.

Our results depend on a notion of symmetrizability defined as follows.

Definition 8 (Symmetrizing distributions)**.** Consider $(V_{\mathbf{s}|\mathbf{x},\mathbf{x}'}, V'_{\mathbf{s}|\mathbf{x},\mathbf{x}'}) \in \Delta(\mathcal{S}|\mathcal{X}^2)^2$. Define $\mathcal{V}$ to be the set of all $(V_{\mathbf{s}|\mathbf{x},\mathbf{x}}, V'_{\mathbf{s}|\mathbf{x},\mathbf{x}'})$ such that for all $(x,x',y) \in \mathcal{X}^2 \times \mathcal{Y}$ it holds that

$$\sum_{s\in\mathcal{S}} V_{\mathbf{s}|\mathbf{x},\mathbf{x}'}(s|x,x') W_{\mathbf{y}|\mathbf{x},\mathbf{s}}(y|x,s) = \sum_{s\in\mathcal{S}} V'_{\mathbf{s}|\mathbf{x},\mathbf{x}'}(s|x',x) W_{\mathbf{y}|\mathbf{x},\mathbf{s}}(y|x',s).$$

B. Main result

Let C be as defined in Equation (3) where $|\mathcal{U}| = K$. In what follows, we show that C is the causal-capacity for AVCs satisfying the assumptions in Section II-B. Equation (3) corresponds to the "babble-and-push" attack outlined below and we also prove a matching achievability. See the proceeding Section IV and [15, Sec. II] for a more comprehensive high-level description of the capacity expression.

$$C := \limsup_{K\to\infty} \max_{\substack{P_{\mathbf{x}|\mathbf{u}}\in\Delta(\mathcal{X}|\mathcal{U}) \\ \left[P_{\mathbf{u}}P_{\mathbf{x}|\mathbf{u}}\right]_{\mathbf{x}}\in\lambda_{\mathbf{x}}}} \min \left\{ \min_{V_{\mathbf{s}|\mathbf{x},\mathbf{u}}\in\mathcal{F}(P_{\mathbf{x}|\mathbf{u}})} I(P_{\mathbf{x}|\mathbf{u}}, V_{\mathbf{s}|\mathbf{x},\mathbf{u}}), \min_{\substack{(\alpha,(V_{\mathbf{s}|\mathbf{x},\mathbf{u}^{\leqslant\alpha}},V_{\mathbf{s}|\mathbf{x},\mathbf{x}',\mathbf{u}^{>\alpha}}))\in\mathcal{A}\times\mathcal{F}_\alpha(P_{\mathbf{x}|\mathbf{u}}) \\ \forall u\in\mathcal{U}^{>\alpha},\,(V_{\mathbf{s}|\mathbf{x},\mathbf{x}',\mathbf{u}=u},V'_{\mathbf{s}|\mathbf{x},\mathbf{x}',\mathbf{u}=u})\in\mathcal{V}}} I(P_{\mathbf{x}|\mathbf{u}}, V_{\mathbf{s}|\mathbf{x},\mathbf{u}^{\leqslant\alpha}}) \right\} \tag{3}$$

Theorem 1 (Converse)**.** *For any $\delta > 0$ rate $C + \delta$ is not achievable for AVCs satisfying the assumptions in Section II-B.*

Theorem 2 (Achievability)**.** *For any $\delta > 0$ rate $C - \delta$ is achievable for AVCs satisfying the assumptions in Section II-B.*

IV. Outline of the Argument

As noted in the introduction, we defer formal proofs to an extended version [15] and outline the main arguments here.

A. Converse

Let $R \geqslant C + \delta$. Let $K = |\mathcal{U}| = \mathrm{poly}_K(1/\delta)$ for a suitable polynomial poly_K (that depends on the parameters of the AVC at hand). James's jamming strategy operates on K chunks each of length n/K. We first describe the strategy when Alice and Bob's code is deterministic. We begin by showing that James can find a subcode which contains a constant fraction of codewords which are chunk-wise approximately constant composition. That is, in each chunk u all codewords in the subcode are approximately typical with respect to some distribution $P_{\mathbf{x}|\mathbf{u}=u}$. It is thus sufficient for James to cause an error on this subcode.

1) James chooses an α representing a threshold point in the code between his "babble" and "push" phases. The former is applied to the first αK codeword chunks and the latter to the remaining chunks.
2) In the "babble" phase James uses channels $V_{\mathbf{s}|\mathbf{x},\mathbf{u}^{\leqslant\alpha}=u}$ for $u = 1, 2, \ldots \alpha K$ and generates $\underline{\mathbf{s}}(t)$ for $t \leqslant \alpha n$ by passing $\underline{x}(t)$ through this channel.
3) James then list-decodes the message based on $(\underline{x}(1), \underline{x}(2), \ldots, \underline{x}(\alpha n))$. He chooses another "spoofing" message from the list and corresponding codeword $\underline{x}'$ (in order to confuse Bob between $\underline{x}$ and $\underline{x}'$).
4) James uses channels $V_{\mathbf{s}|\mathbf{x},\mathbf{x}',\mathbf{u}^{>\alpha}=u}$ for $u = \alpha K + 1, \alpha K + 2, \ldots, K$ in the remaining chunks to generate $\underline{\mathbf{s}}(t)$ by passing the pair $(\underline{x}(t), \underline{x}'(t))$ through the channel $V_{\mathbf{s}|\mathbf{x},\mathbf{x}',\mathbf{u}=u}$ in the u-th chunk.

This strategy must satisfy two conditions. First, it has to satisfy the cost constraint, which is the sum of the cost from each phase. The technical challenge comes in the second phase because the joint type of $\underline{x}$ and $\underline{x}'$ need not be a product type. However, due to the Generalized Plotkin bound [17] we show that the suffix state cost only needs to be computed with respect to distributions $P_{\mathbf{x},\mathbf{x}'}$ that are convex combinations of product distributions. These conditions constrain the induced jamming distributions of James minimized over in the capacity expression above (Equation (3)).

The second condition is that the "push" channels must be symmetrizing (see Definition 8) ensuring that the suffix $(\underline{y}(\alpha n + 1), \ldots, \underline{y}(n))$ observed by Bob was equally likely to have been generated by either the codeword suffix $\underline{x}^{>\alpha} = (\underline{x}(\alpha n + 1), \ldots, \underline{x}(n))$ or the codeword suffix $\underline{x}'^{>\alpha} = (\underline{x}'(\alpha n + 1), \ldots, \underline{x}'(n))$. This forces a constant probability of error for Bob since he cannot distinguish the true message/codeword from the spoofing message/codeword chosen by James.

To show that this strategy will work, we show via relatively standard information-theoretic strong converse arguments that for rate R exceeding the capacity expression above, if James chooses a break point α such that the cumulative mutual information thus far is "just" below the rate, then Bob must still have a large set (say of size $2^{\Omega(n)}$) of codewords consistent with what he has received. Here we use the assumption that the channel is state deterministic: *James can also compute the same list as Bob.* Then, via analysis paralleling the symbol-error analysis [10], but significantly generalizing it to a general class of AVCs, we show that the push strategy outlined above will result in a constant probability (that depends on δ) of James being able to cause a decoding error by Bob.

As mentioned, a key ingredient in the analysis is the Generalized Plotkin bound [17] which ensures that with positive probability the joint type of randomly sampled pairs of codewords (more precisely, their suffixes) is a convex combination of product distributions. This results in significant cost savings for James. To guarantee James' success (with constant probability) in the setting of stochastic encoders as well, we use information-theoretic and coding-theoretic techniques introduced in prior work [11].

B. Achievability

Let $R \leqslant C - \delta$. Alice uses a chunk-wise constant composition stochastic code with distributions $\{P_{\mathbf{x}|\mathbf{u}=u}\}_{u=1}^K$ such that the average composition $P_{\mathbf{x}} = \frac{1}{K}\sum_{u=1}^K P_{\mathbf{x}|\mathbf{u}=u}$ satisfies the input constraint. She uses $K = \mathrm{poly}_K(1/\delta)$ chunks and uses $\mathrm{poly}_r(\delta)n$ bits of private randomness, for suitable polynomials poly_K and poly_r (that, again, depend on the parameters of the AVC at hand). For every message m, chunk u, and randomness r, Alice selects codeword-chunks of length n/K uniformly from type $P_{\mathbf{x}|\mathbf{u}=u}$. This use of encoder randomness, which is independent in each chunk, forces uncertainty for James about the actual codeword that will be transmitted in the next chunk when conditioning on prior chunks. Such randomly designed codes have several nice properties used in our analysis. In particular, using AVC list-decoding arguments, adapted to the chunkwise stochastic encoding scheme here,

our codes are (w.h.p.) list-decodable with list size $\mathrm{poly}(1/\delta)$. In addition, extending concentration measures analyzed in [10], codewords $\underline{x}'$ in the lists obtained through list-decoding have (with high probability over Alice's private randomness) joint type approximately $P_{\mathbf{x},\mathbf{x}'|\mathbf{u}=u} = P^{\otimes 2}_{\mathbf{x}|\mathbf{u}=u}$ with the transmitted codeword $\underline{x}$ (for every chunk index u).

Bob's decoding process is an iterative process reminiscent of the 2-phase process of James presented in the converse proof. Without any prior knowledge of James' jamming strategy $\underline{s}$, Bob iterates over potentially transmitted codewords $\underline{x}'$ and state vectors $\underline{s}'$ that may have resulted with the received word $\underline{y}$. More specifically, taking causality into account, Bob iterates over all potential jamming strategies $V'_{\mathbf{s}|\mathbf{x},\mathbf{u}}$ that satisfy the jamming cost constraint with respect to $P_{\mathbf{x}|\mathbf{u}}$ and for each such V' proceeds in 2-phases. First, he chooses a threshold α' corresponding to V' such that $\alpha' K$ is the first chunk for which the normalized cumulative mutual information (Definition 7) is *more* than the rate R of the code, and list decodes under the (potentially mistaken) assumption that James is acting according to $V'_{\mathbf{s}|\mathbf{x},\mathbf{u}^{\leqslant\alpha'}}$. As mentioned above, Bob's obtained list will be of size $\mathrm{poly}(1/\delta)$. Notice that, depending on the studied V', the transmitted message m and the corresponding codeword $\underline{x}$ may or may not be in the list. Bob's objective at this point is two-fold: to reject the list if it does not include the transmitted codeword and, otherwise, to disambiguate the list, i.e., find $\underline{x}$. Both objectives are accomplished in the second phase of Bob's decoding.

In the second decoding phase, Bob examines each codeword $\underline{x}'$ in his obtained list and outputs the message m' corresponding to $\underline{x}'$ if and only if $\underline{y}$ could have been obtained from $\underline{x}'$ through a feasible $\underline{s}'$. In particular, if there exists a vector $\underline{s}'$ and a corresponding $(V'_{\mathbf{s}|\mathbf{x},\mathbf{u}^{\leqslant\alpha'}}, V'_{\mathbf{s}|\mathbf{x},\mathbf{x}',\mathbf{u}^{>\alpha'}}) \in \mathcal{F}_{\alpha'}(P_{\mathbf{x}|\mathbf{u}})$ representing the transition from $\underline{x}'$ to $\underline{y}$ through $\underline{s}'$. Here, $(V'_{\mathbf{s}|\mathbf{x},\mathbf{u}^{\leqslant\alpha'}}, V'_{\mathbf{s}|\mathbf{x},\mathbf{x}',\mathbf{u}^{>\alpha'}})$ is a refinement of $V'_{\mathbf{s}|\mathbf{x},\mathbf{u}}$ derived using the joint type of $\underline{x}'$ and the assumed $\underline{s}'$ (both known to Bob), and the transmitted $\underline{x}$ (unknown to Bob) where $(\underline{x},\underline{x}')$ are assumed through code design to have type $P^{\otimes 2}_{\mathbf{x}|\mathbf{u}=u}$. If no $\underline{x}'$ in the list passes the test above, Bob continues in studying the next jamming strategy $V'_{\mathbf{s}|\mathbf{x},\mathbf{u}}$ until eventually finding a codeword and a corresponding message that pass the test. Once such a codeword is found the decoding process is terminated. If no codewords pass the test in any of Bob's iterations, a decoding error is considered.

In our analysis, we show, for every transmitted message m, that with high probability over the stochasticity of Alice, the decoding process will succeed. Namely, with high probability over the stochasticity of Alice, only the codeword $\underline{x}$ corresponding to Alice's message m will pass the decoding test. The proof involves a careful ordering on potential jamming strategies $V'_{\mathbf{s}|\mathbf{x},\mathbf{u}}$ used by Bob in the decoding process and the corresponding analysis for general channels requires more care than the specific models considered in prior work. Specifically, we order V' based on their corresponding thresholds α' (from small to large). For such an ordering, it suffices to show (i) that once Bob studies the *correct* $V_{\mathbf{s}|\mathbf{x},\mathbf{u}}$ corresponding to the actual (unknown to Bob) jamming vector $\underline{s}$ (or more precisely, the correct triple $(\alpha, (V_{\mathbf{s}|\mathbf{x},\mathbf{u}^{\leqslant\alpha}}, V_{\mathbf{s}|\mathbf{x},\mathbf{x}',\mathbf{u}^{>\alpha}}))$), then using standard typicality arguments the message m will pass the test of the decoder; and (ii) that for any "weaker" jammer V', i.e., with $\alpha' \leqslant \alpha$, where α corresponds to the actual jamming strategy of James through V above, no message will pass the decoding test.

The main technical difficulty in our proof addresses case (ii) above. Suppose, to the contrary, that Bob can decode to an incorrect codeword $\underline{x}'$ using the "wrong" choice of $(\alpha', (V'_{\mathbf{s}|\mathbf{x},\mathbf{u}^{\leqslant\alpha'}}, V'_{\mathbf{s}|\mathbf{x},\mathbf{x}',\mathbf{u}^{>\alpha'}}))$ with $\alpha' \leqslant \alpha$, implied by a state vector $\underline{s}'$ which can generate $\underline{y}$ from $\underline{x}'$. Considering the true jamming vector $\underline{s}$ of James, and its corresponding single letter representation $(\alpha, (V_{\mathbf{s}|\mathbf{x},\mathbf{u}^{\leqslant\alpha}}, V_{\mathbf{s}|\mathbf{x},\mathbf{x}',\mathbf{u}^{>\alpha}}))$, which can generate $\underline{y}$ from $\underline{x}$ and $\underline{s}$, we have that the joint type of $(\underline{x},\underline{x}',\underline{y})$ can be obtained in two different ways: through $\underline{x}'$ and $\underline{s}'$ via V', and through $\underline{x}$ and $\underline{s}$ via V, implying that V' and V are a symmetrizing pair. More precisely, to reach a contradiction to the definition of C in (3), and thus conclude our proof, we must show that $(V'_{\mathbf{s}|\mathbf{x},\mathbf{x}',\mathbf{u}^{>\alpha'}=u}, V_{\mathbf{s}|\mathbf{x},\mathbf{x}',\mathbf{u}^{>\alpha'}=u})$ are a symmetrizing pair for every chunk $u = \alpha K + 1, \alpha K + 2, \cdots, K$, and that (without loss of generality) $V'_{\mathbf{s}|\mathbf{x},\mathbf{u}^{\leqslant\alpha'}} = V_{\mathbf{s}|\mathbf{x},\mathbf{u}^{\leqslant\alpha'}}$. For the latter, one requires a crucial and novel step in our analysis that relies on our iterative ordering of V' and in particular on the fact that $\alpha' \leqslant \alpha$. Assumptions 3 and 5 on our AVC appearing in Section II-B are distilled from this step in our analysis. The former is then proven through a generalization of the analysis for oblivious AVCs [4].

V. Discussion

In this paper we characterized the capacity for AVCs with online adversaries when the channel is state-deterministic. This restriction on the channel class is used only in the converse, where James must be able to compute the channel output exactly, similar to previous "snooping" models for bit flips/erasures [14]. It may be possible to extend this argument to more general classes of AVCs, which we leave for future work. On the achievability side, we note that perhaps surprisingly, input distributions $P_{\mathbf{x}|\mathbf{u}=u}$ that are i.i.d. (i.e. there is some $P_{\mathbf{x}}$ such that $P_{\mathbf{x}|\mathbf{u}=u} = P_{\mathbf{x}}$ for all u) do not necessarily attain capacity for general AVCs. This is consistent with prior work: while i.i.d. input distributions attain capacity for finite alphabet symbol-error/erasure channels with no input constraints [10], they do not for the quadratically constrained casual model [12]. This seems counterintuitive given the convexity of mutual information function and the linearity of the symmetrizability condition. The reason is that Alice has to choose a rate R and design her corresponding codebook distribution to be simultaneously good for any possible choice of α (and the corresponding feasible input distributions), rendering the problem highly non-convex. This reveals another interesting question for future work: are there computationally tractable approximations for our capacity expression that will reveal optimal encoder/jamming strategies?

REFERENCES

[1] C. Shannon, "A Mathematical Theory of Communication," *Bell System Technical Journal*, vol. 27, pp. 379–423, 623–656, 1948. I

[2] D. Blackwell, L. Breiman, and A. J. Thomasian, "The Capacity of a Class of Channels under Random Coding," *Ann. of Mathematical Statistics*, vol. 31, no. 3, pp. 558–567, 1960. I

[3] I. Csiszár and P. Narayan, "Arbitrarily varying channels with constrained inputs and states," *IEEE Trans. Inf. Theory*, vol. 34, pp. 27–34, 1988. I, II-A

[4] ——, "The Capacity of the Arbitrarily Varying Channel Revisited : Positivity, Constraints," *IEEE Trans. Inf. Theory*, vol. 34, pp. 181–193, 1988. I, III, IV-B

[5] M. Langberg, S. Jaggi, and B. K. Dey, "Binary causal-adversary channels," in *International Symposium on Information Theory (ISIT)*, Seoul, South Korea, June 28–July 3 2009, pp. 2723–2727. I

[6] B. K. Dey, S. Jaggi, M. Langberg, and A. D. Sarwate, "Coding against delayed adversaries," in *International Symposium on Information Theory (ISIT)*, 2010, pp. 285–289. I

[7] B. Dey, S. Jaggi, and M. Langberg, "Codes Against Online Adversaries: Large Alphabets," *IEEE Trans. Inf. Theory*, vol. 59, pp. 3304–3316, June 2013. I

[8] R. Bassily and A. Smith, "Causal Erasure Channels," in *Proc. ACM Symp. on Discrete Algorithms (SODA)*, Hong Kong, China, January 2014. I

[9] Z. Chen, S. Jaggi, and M. Langberg, "A characterization of the capacity of online (causal) binary channels," in *Proceedings of the forty-seventh annual ACM symposium on Theory of computing*, 2015, pp. 287–296. I

[10] ——, "The capacity of online (causal) q-ary error-erasure channels," *IEEE Transactions on Information Theory*, vol. 65, no. 6, pp. 3384–3411, 2019. I, IV-A, IV-B, V

[11] B. Dey, S. Jaggi, M. Langberg, and A. Sarwate, "Upper Bounds on the Capacity of Binary Channels With Causal Adversaries," *IEEE Trans. Inf. Theory*, vol. 59, pp. 3753–3763, June 2013. I, IV-A

[12] T. Li, B. K. Dey, S. Jaggi, M. Langberg, and A. D. Sarwate, "Quadratically constrained channels with causal adversaries," in *International Symposium on Information Theory (ISIT)*. IEEE, 2018, pp. 621–625. I, 1, V

[13] A. Mazumdar, "On the capacity of memoryless adversary," in *International Symposium on Information Theory (ISIT)*, 2014, pp. 2869–2873. I

[14] V. Suresh, E. Ruzomberka, and D. J. Love, "Stochastic-adversarial channels: Online adversaries with feedback snooping," in *2021 IEEE International Symposium on Information Theory (ISIT)*, 2021, pp. 497–502. I, V

[15] Y. Zhang, S. Jaggi, M. Langberg, and A. D. Sarwate, "The capacity of causal adversarial channels," arXiv preprint arXiv:2205.06708 (2022). I, 3, 5, III-B, IV

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

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