

Channels with nosy “noise”

Anand D. Sarwate and Michael Gastpar

Department of Electrical Engineering and Computer Sciences

University of California, Berkeley

Berkeley, CA 94720, USA

Email: {asarwate, gastpar}@eecs.berkeley.edu

Abstract—Coding over channels whose state can depend non-causally on the entire transmitted codeword and message are studied. The channel model is a variation on the arbitrarily varying channel (AVC) with state constraints. The randomized coding capacity of this channel is shown to be equal to the minimum of the capacities of channels in the row-convex closure of the AVC. Common randomness of $O(\log n)$ bits is sufficient to achieve this capacity.

I. INTRODUCTION

In studying the problem of communication in the presence of an adversary or *jammer*, the capabilities and knowledge of the jammer are of primary importance in specifying the model. The jammer may have no knowledge of the message to be transmitted, may know the message but not the transmitted codeword, may be able to wiretap the channel to get causal knowledge of the codeword, or may know the full codeword in advance. Coding strategies under one set of assumptions on the jammer may no longer be effective under another. In the arbitrarily varying channel (AVC) many of these distinctions can be captured in the definition of the error probability. For example, in the for deterministic codes under maximum error, it is implicit that the jammer may choose its input based on the entire transmitted codeword. For randomized codes, the jammer is assumed to know the *message* but not the *codeword*. In this work we will study randomized coding for the case in which the jammer knows *both the message and the entire codeword*, as shown in Figure 1.

In a randomized code, the encoder and decoder share a source of common randomness or *secret key* that is unknown to the jammer. This introduces randomness in the mapping from a message to its codeword. Randomized codes are typically designed to make the jammer’s actions no worse than noise from the perspective of the decoder. The assumption in standard AVCs under randomized coding is that the jammer may know the message but not the codeword. In the case where the jammer knows the codeword as well, more care must be taken in the code construction, since a particular codeword may give the jammer enough information about the secret key to render the randomization useless.

Coding for AVCs with partial or full knowledge of the codeword at the jammer has been studied before. The classic paper of Blackwell, Breiman, and Thomasian [1] proves a randomized coding theorem when the jammer can observe the channel inputs and outputs causally, and related questions are discussed in the book by Csiszár and Körner [2]. For

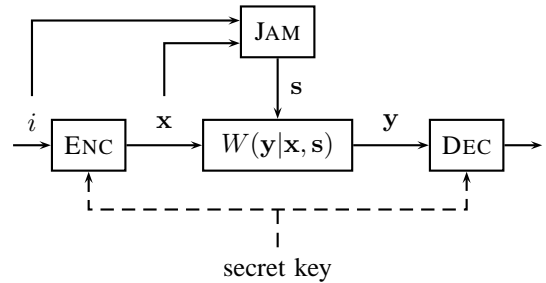


Fig. 1. Communicating over a channel whose state can depend on the message i and channel input \mathbf{x} non-causally. The jammer can choose \mathbf{s} to be a function of i and \mathbf{x} but does not have access to the common randomness (secret key) shared by the encoder and decoder.

this model, deterministic coding was also investigated by Ahlswede and Wolfowitz [3], but in general the deterministic coding capacity is unknown. The capacity under randomized coding for binary channels with a constrained bit-flipping jammer that knows the codeword was recently solved by Langberg [4], who showed that if the jammer can flip no more than a fraction Λ of the bits, the capacity is $C(\Lambda) = 1 - h_b(\Lambda)$, the same as a binary symmetric channel with flip probability Λ . Furthermore, he showed that $O(\log n)$ bits of common randomness were sufficient for codes of blocklength n . Smith [5] showed that computationally efficient capacity-achieving codes exist for this channel with $O(n)$ bits of common randomness. Agarwal, Sahai, and Mitter [6] used randomized codes for an adversarial channel model in which the adversary is restricted by a distortion constraint and whose capacity is a rate-distortion function. In our model, the structure of the channel and constraints are different, and we do not fix the input distribution.

In this work we extend the result of [4] to general cost-constrained AVCs in which the codeword is known to the jammer. The capacities for randomized coding under maximal error $C_r(\Lambda)$ and deterministic coding under average error with no codeword information at the jammer were found by Csiszár and Narayan [7], [8]. In their randomized code construction, the amount of shared randomness between the encoder and decoder is not bounded explicitly and it is assumed that the jammer does not know the transmitted codeword. Unlike in the binary additive case, we cannot achieve the randomized coding capacity $C_r(\Lambda)$. Instead, the capacity is limited by the

worst case average channel with input-dependent state:

$$\bar{\mathcal{W}}_{\text{dep}}(\Lambda) = \left\{ V(y|x) : V(y|x) = \sum_s W(y|x, s)U(s|x) \right\}$$

The corresponding capacity is:

$$C^{\text{dep}}(\Lambda) = \max_{P(x)} \min_{V \in \bar{\mathcal{W}}_{\text{dep}}(\Lambda)} I(X \wedge Y) \quad (1)$$

The central result of this paper is that the randomized coding capacity can be achieved with $O(\log n)$ bits of common randomness even when the jammer knows the codeword.

Main Theorem 1: For the AVC $\{\mathcal{W}, l(s), \Lambda\}$ with codeword known to the jammer, the rate $R(n)$ is achievable with a secret key of $\log K(n)$ bits and error $\hat{\epsilon}_r(n)$, where

$$R(n) = C^{\text{dep}}(\Lambda) - \rho(n) - \frac{1}{2n} \log K(n) \quad (2)$$

$$\rho(n) \leq 4C^{\text{dep}}(\Lambda) \log |\mathcal{Y}| \cdot \frac{n}{\hat{\epsilon}_r(n) \sqrt{K(n)} \log K(n)} \quad (3)$$

One strategy for constructing randomized codes with small secret key comes from the "elimination technique" introduced by Ahlswede [9], which involves sampling a randomized code with large secret key to get a code with smaller key size whose probability of decoding error can still be driven to 0. This technique has also been used to bound the randomization needed when the jammer does not know the codeword [10]. When the jammer does know the transmitted codeword, such a sampling procedure will not work because the jammer's knowledge will introduce dependencies in the error.

Instead, we use the approach suggested by Langberg [4], which uses deterministic list codes followed by a derandomization step. We first generalize the results of Ahlswede [11], [12] to the constrained AVC setting and show that a rate $R = C^{\text{dep}}(\Lambda) - \epsilon$ can be achieved with list size $L = O(\epsilon^{-1})$ under maximum error. We then describe a way of constructing a randomized code by taking subsets of the list code's codewords. In the next section we describe the problem more formally, and then describe our results on list codes in Section III. In Section IV we provide the randomized code construction, and we discuss some of the applications of this result in Section V.

II. DEFINITIONS

An arbitrarily varying channel (AVC) is a collection of $\mathcal{W} = \{W(\cdot|\cdot, s) : s \in \mathcal{S}\}$ of channels from an input alphabet \mathcal{X} to an output alphabet \mathcal{Y} parameterized by a state $s \in \mathcal{S}$. Here we will assume the sets \mathcal{X} , \mathcal{Y} and \mathcal{S} are finite. If $\mathbf{x} = (x_1, x_2, \dots, x_n)$, $\mathbf{y} = (y_1, y_2, \dots, y_n)$ and $\mathbf{s} = (s_1, s_2, \dots, s_n)$ are length n vectors, the probability of \mathbf{y} given \mathbf{x} and \mathbf{s} is given by:

$$W(\mathbf{y}|\mathbf{x}, \mathbf{s}) = \prod_{i=1}^n W(y_i|x_i, s_i) \quad (4)$$

We think of the state as being controlled by a malicious adversary, called the *jammer*, whose objective is to maximize the probability of decoding error and thereby minimize the capacity.

We are interested in the case where there is a cost function $l : \mathcal{S} \rightarrow \mathbb{R}^+$ on the jammer. We will assume that $\max_s l(s) \leq l_{\max} < \infty$. The cost of an n -tuple is

$$l(\mathbf{s}) = \sum_{k=1}^n l(s_k) \quad (5)$$

The state obeys a state constraint Λ if

$$l(\mathbf{s}) \leq n\Lambda \quad a.s. \quad (6)$$

Note that if $\Lambda = l_{\max}$ then the state constraint is inoperative and we return to the unconstrained case.

As an example, take the AVC given by $y = x \oplus s$ with $\mathcal{X} = \mathcal{Y} = \mathcal{S} = \{0, 1\}$ and $l(s) = s$. This is similar to a binary symmetric channel (BSC) with flip probability Λ . Here, however, the channel is constrained to flip *no more than* Λn bits but a code must correct *any error pattern* of weight Λn . By contrast, a BSC will flip close to Λn bits with *high probability* and the code must correct *most error patterns*.

Let $[M] = \{1, 2, \dots, M\}$. An (n, N, K) randomized code for the AVC is a family of maps $\{(\phi_k, \psi_k) : k = 1, 2, \dots, K\}$ indexed by a set of K keys. The encoding maps are functions $\phi_k : [N] \rightarrow \mathcal{X}^n$ and the decoding maps are $\psi_k : \mathcal{Y}^n \rightarrow [N]$. The rate of the code is $R = n^{-1} \log N$. The decoding region for message i under key k is $D_{i,k} = \{\mathbf{y} : \psi_k(\mathbf{y}) = i\}$. In the case where the jammer knows the codeword we will want the images of the encoding maps $\text{im } \phi_k$ to have large intersection so that the knowledge of the codeword does not reveal too much about the key k .

The power of randomized codes comes from modifying the definition of the error probability. Rather than demanding that the decoder error be small for every message and every key value, we instead require it to be small for every message on *average over key values*. Here we assume the key is chosen uniformly in the set $\{1, 2, \dots, K\}$. For standard AVC in which the jammer does not know the codeword, we define the maximum probability of error by

$$\epsilon_r = \max_i \max_s \frac{1}{K} \sum_{k=1}^K (1 - W^n(D_{i,k}|\phi_k(i), \mathbf{s})) \quad (7)$$

Let $\mathcal{J}(\Lambda) = \{J : [N] \times \mathcal{X}^n \rightarrow \mathcal{S}^n : l(J(\mathbf{x})) \leq n\Lambda\}$. When the jammer knows both the message i and the transmitted codeword $\phi_k(i)$, then we define the maximum probability of error by

$$\hat{\epsilon}_r = \max_i \max_{J \in \mathcal{J}(\Lambda)} \frac{1}{K} \sum_{k=1}^K (1 - W^n(D_{i,k}|\phi_k(i), J(i, \phi_k(i)))) \quad (8)$$

We will say a sequence of pairs $(R(n), K(n))$ is *achievable under maximum error* if there exists a sequence of $(n, 2^{nR(n)}, K(n))$ randomized codes whose error $\epsilon \rightarrow 0$ as $n \rightarrow \infty$. The capacity with randomization $K(n)$ is the supremum of all achievable rates.

An (n, N, L) deterministic list code for the AVC is a pair of maps (ϕ, ψ) where the encoding function is ϕ :

$\{1, 2, \dots, N\} \rightarrow \mathcal{X}^n$ and the decoding function is $\psi : \mathcal{Y}^n \rightarrow \{1, 2, \dots, N\}^L$. The rate of the code is $R = n^{-1} \log(N/L)$. The codebook is the set of vectors $\{\mathbf{x}_i : 1 \leq i \leq N\}$, where $\mathbf{x}_i = \phi(i)$. The decoding region for message i is $D_i = \{\mathbf{y} : i \in \psi(\mathbf{y})\}$. We will often specify a code by the pairs $\{(\mathbf{x}_i, D_i) : i = 1, 2, \dots, N\}$, with the encoder and decoder implicitly defined.

For list codes we can also define the maximum probability of error:

$$\varepsilon_L = \max_i \max_s (1 - W^n(D_i | \phi(i), \mathbf{s})) . \quad (9)$$

Because the decoding regions D_i can overlap for list codes, the error probability can be small without randomization for every message i . A sequence $(R(n), L(n))$ is achievable if there exists a sequence of $(n, 2^{nLR(n)}, L(n))$ list codes whose error $\varepsilon_L \rightarrow 0$ as $n \rightarrow \infty$. The list coding capacity for list codes of list size L is the supremum of rates achievable with fixed list size $L(n) = L$ and is denoted by $C_L(\Lambda)$.

III. LIST CODES FOR STATE-CONSTRAINED AVCs

The arbitrarily varying channel with deterministic codes and maximal error is directly related to the design of error correcting codes. Because this is a difficult problem, we can instead consider list decoding, a relaxation in which the decoder is allowed to output a small list and we need only guarantee the transmitted message is in the list.

Theorem 1 (List decoding for maximal error): Let $\mathcal{W} = \{W(\cdot | \cdot, s) : s \in \mathcal{S}\}$ be an arbitrarily varying channel with constraint function $l(s)$ and state constraint Λ . Fix a rate $R < C^{\text{dep}}(\Lambda)$. Then R is achievable under maximal error with deterministic list codes of list size

$$L < O\left(\frac{1}{C^{\text{dep}}(\Lambda) - R}\right) . \quad (10)$$

In other words,

$$C_d(L, \Lambda) \geq C^{\text{dep}}(\Lambda) - O(L^{-1}) . \quad (11)$$

The result is proved in two steps – first we claim that list codes of exponential list size exist, and then we construct a code of finite list size by sampling codewords from the larger list code. This line of argument follows that developed by Ahlswede [11], [12].

A. List codes with large lists

Lemma 1: Let $(\mathcal{W}, l(\cdot), \Lambda)$ be a constrained AVC. For any $\epsilon > 0$ there is an n sufficiently large and an (n, N, L) list code \mathcal{C} with

$$N \geq \exp(n(H(P(x)) - o(\epsilon))) \quad (12)$$

$$L \leq \exp\left(n\left(\max_{V \in \mathcal{W}_{\text{dep}}(\Lambda)} H(V'(x|y) | P'(y)) + O(\epsilon \log \epsilon^{-1})\right)\right) \quad (13)$$

$$\varepsilon(\mathcal{C}) \leq \exp(-nE(\epsilon)) \quad (14)$$

where $P(x)V(y|x) = P'(y)V'(x|y)$.

Details of the proof can be found in [13]. The arguments are similar to [11] and use type arguments of a standard nature [2].

B. List reduction

We can now prove a generalization of Ahlswede's result [12] to the constrained AVC. Given a small gap ϵ from capacity, we subsample the previous code with exponential list sizes to obtain a code with finite list size $O(\epsilon^{-1})$ that can achieve rates ϵ away from capacity.

Lemma 2: Let $(\mathcal{W}, l(\cdot), \Lambda)$ be a constrained AVC whose randomized coding capacity is $C^{\text{dep}}(\Lambda)$. For any $\epsilon' > 0$ there exists a list code of rate $R > C^{\text{dep}}(\Lambda) - \epsilon'$ and list size

$$L' < \left\lceil \frac{4 \log |\mathcal{Y}|}{C^{\text{dep}}(\Lambda) - R} \right\rceil + 1. \quad (15)$$

Proof: By Lemma 1 there exists N, L , and δ satisfying (12) – (14) for any ϵ so that there exists an (n, N, L) list code $\mathcal{C}_L = \{(\mathbf{u}_i, D_i) : i \in [N]\}$. Note that $N/L = \exp(n(C^{\text{dep}} - 2\epsilon))$. We will subsample this codebook to find our code of constant list size.

Let $N' = \exp(nR)$ and $\mathcal{C}_{L'} = \{(\mathbf{x}_j, D_j) : j \in [N']\}$ be a collection of N' codewords selected uniformly from \mathcal{C}_L . We will prove that there exists a constant L' such that no $\mathbf{y} \in \mathcal{Y}^n$ is in more than L' decoding sets D_j with high probability. Fix $\mathbf{y} \in \mathcal{Y}^n$ and note that from the definition of \mathcal{C}_L we have

$$\mathbb{E}(\mathbf{1}(\mathbf{y} \in D_j)) = \mathbb{P}(\mathbf{y} \in D_j) \leq \frac{L}{N} . \quad (16)$$

For a fixed \mathbf{y} , the chance that more than L' decoding regions contain \mathbf{y} out of N' choices can be bounded above using Sanov's Theorem [14, Theorem 12.4.1]. That is, for $\mu = (N' + 1)^2 < \exp(n3R) \leq \exp(n3 \log |\mathcal{Y}|)$ we can choose n large enough so that

$$\mathbb{P}\left(\frac{1}{N'} \sum_{j=1}^{N'} \mathbf{1}(\mathbf{y} \in D_j) > \frac{L'}{N'}\right) \leq \mu \cdot \exp\left(-N' \left(D\left(\frac{L'}{N'} \parallel \frac{L}{N}\right)\right)\right) . \quad (17)$$

We will upper bound the exponent:

$$\begin{aligned} & -N' D\left(\frac{L'}{N'} \parallel \frac{L}{N}\right) + \log \mu \\ &= -L' \log \frac{L'/N'}{L/N} - N' \left(1 - \frac{L'}{N'}\right) \log \frac{1 - L'/N'}{1 - L/N} + \log \mu. \end{aligned} \quad (18)$$

To deal with the second term we use the inequality $-(1-a) \log(1-a) \leq 2a$ (for small a) on the term $(1 - L'/N') \log(1 - L'/N')$ and discard the small positive term $-(1 - L'/N') \log(1 - L/N)$.

$$\begin{aligned} & -N' D\left(\frac{L'}{N'} \parallel \frac{L}{N}\right) + \log \mu \\ & \leq -L' \log \frac{L'/N'}{L/N} + N' \left(2 \frac{L'}{N'}\right) + \log \mu \\ & = -nL'(C^{\text{dep}} - R - 2\epsilon) - L' \log L' + 2L' + \log \mu \\ & \leq -nL'(C^{\text{dep}} - R - 2\epsilon) + 3 \log |\mathcal{Y}| \end{aligned} \quad (19)$$

← message →				
\sqrt{K}	\mathbf{x}_3, A_{11}	\mathbf{x}_5, A_{12}	\mathbf{x}_{16}, A_{13}	• • •
	\mathbf{x}_{81}, A_{21}	\mathbf{x}_9, A_{22}	\mathbf{x}_{22}, A_{23}	• • •
	\mathbf{x}_4, A_{31}	\mathbf{x}_{63}, A_{32}	\mathbf{x}_2, A_{33}	• • •
	\vdots	\vdots	\vdots	
	$\mathbf{x}_1, A_{\sqrt{K}1}$	$\mathbf{x}_7, A_{\sqrt{K}2}$	$\mathbf{x}_7, A_{\sqrt{K}3}$	• • •
\sqrt{K}				

Fig. 2. Constructing a randomized code from a list-decodable code. We put the codewords of the list code into a $\sqrt{K} \times N/\sqrt{K}$ table. Each column has a partition of the set of K keys into sets A_{ij} of \sqrt{K} keys each. The intersection of the key sets is small.

The last inequality follows from taking $L' > 4$ and the bound on μ .

Now we take a union bound over all \mathbf{y} to get

$$\mathbb{P} \left(\frac{1}{N'} \sum_{j=1}^{N'} \mathbf{1}(\mathbf{y} \in D_j) > \frac{L'}{N'} \quad \forall \mathbf{y} \right) \leq \exp(-n(L'(C^{\text{dep}} - R - 2\epsilon) + 4 \log |\mathcal{Y}|)) . \quad (20)$$

Then we have for

$$L' > \left\lfloor \frac{4 \log |\mathcal{Y}|}{C^{\text{dep}} - R - 2\epsilon} \right\rfloor + 1 , \quad (21)$$

have an $(n, \exp(nR), L')$ codebook. Letting $\epsilon' = 3\epsilon$ and $R = C^{\text{dep}} - \epsilon'$, we have a list code with lists of size $O(1/\epsilon')$, as desired. ■

Theorem 1 now follows from the preceding Lemma.

IV. FROM LIST CODES TO RANDOMIZED CODES

Given a list code of small (constant) list size, we can use a small amount of common randomness to construct a randomized code by using a type of cryptographic message authentication system. This construction has been used by Langberg [4] and Smith [5] to construct randomized codes for constrained bit-flipping AVCs in which the codeword is known to the jammer. By using our new list codes we can construct such randomized codes for general AVCs. In this section we briefly describe the scheme and use out new results on list codes to characterize the error probability with the key size and target rate.

Lemma 3: Let \mathcal{C}_L be an (n, N, L) deterministic list code of rate $R = n^{-1} \log N$ and probability of error δ . For key size $K(n)$ there exists an $(n, N/\sqrt{K(n)}, K(n), \delta + \delta')$ randomized code where

$$\delta' = \frac{2LnR}{\sqrt{K} \log K} . \quad (22)$$

Proof: Let $R' = n^{-1} \log(N/\sqrt{K})$. Let i and z be elements of $GF(\sqrt{K})$, and let the key be given by the pair (i, z) . Pick $\{f_j(\cdot) : j = 1, 2, \dots, 2^{nR'}\}$ to be a set of $2^{nR'}$ distinct monic polynomials of degree $d - 1$ over $GF(\sqrt{K})$. Then let $A_{ij} = \{(i, z, f_j(z) + i) : z = 1, 2, \dots, \sqrt{K}\}$.

Since i is a constant shift of the polynomial $f_j(\cdot)$, it is clear that $\{A_{ij} : i = 1, 2, \dots, \sqrt{K}\}$ is a partition of the set of all keys. Furthermore, for $j' \neq j$ we have $|A_{ij} \cap A_{ij'}| \leq d$, since $f_j(z) = f_{j'}(z)$ for at most d values of z .

We now construct a table as shown in Figure 2. The columns index the N/\sqrt{K} messages for the randomized code, and the rows the \sqrt{K} possible values of i . The N codewords $\{\mathbf{x}_l\}$ of the list code \mathcal{C}_L are randomly placed in the table. We also associate A_{ij} with the (i, j) -th cell in the table. The encoder takes a message j and key (i, z) and outputs the codeword of the list code in the (i, j) -th position in the table. Note that knowledge of the transmitted codeword tells the jammer both the message j and part of the key i .

The decoder for the randomized code first decodes using the list code \mathcal{C}_L to find a list of at most L candidate codewords $\{\mathbf{x}_{l_1}, \mathbf{x}_{l_2}, \dots, \mathbf{x}_{l_L}\}$. Each of these codewords has an associated key set $\{A_{ij}(l_1), \dots, A_{ij}(l_L)\}$ given by the table. The decoder chooses the unique l_k for which $(i, z) \in A_{ij}(l_k)$ (if it exists) and outputs the corresponding message j_k .

There are two possible decoding errors. If the list code has a decoding error then the correct codeword will not be in the list and so the decoder for the randomized code will fail. This happens with probability smaller than δ by the assumptions on the list code. If the transmitted codeword is in the list produced by the list decoder, then we will have an error if there is not a unique l_k for which $(i, z) \in A_{ij}(l_k)$. That is, we must have $(i, z) \in A_{ij_{k'}}$ for some $k' \neq k$. We know $|A_{ij} \cap A_{ij'}| \leq d$, so there are at most Ld values of (i, z) for which this can happen. Since the jammer knows i and there are \sqrt{K} values for z , the probability that the key cannot disambiguate the list is at most $\delta' = Ld/\sqrt{K}$. The total error probability is then $\delta + \delta'$.

The last part is to choose d appropriately. There are \sqrt{K}^{d-1} monic polynomials of degree $d-1$ over $GF(\sqrt{K})$, so we need

$$\sqrt{K}^{d-1} \geq \frac{1}{\sqrt{K}} 2^{nR} . \quad (23)$$

This in turn implies

$$d \geq \frac{nR}{\log \sqrt{K}} . \quad (24)$$

Substituting this into the expression for δ' in the previous paragraph we obtain (22) ■

Theorem 2: For the AVC $\{\mathcal{W}, l(s), \Lambda\}$ with codeword known to the jammer, the pairs $(R'(n), K(n))$ are achievable with error $\hat{\epsilon}_r(n)$, where

$$R'(n) = C^{\text{dep}}(\Lambda) - \rho(n) - \frac{1}{2n} \log K(n) \quad (25)$$

$$\rho(n) \leq 16C^{\text{dep}}(\Lambda) \log |\mathcal{Y}| \cdot \frac{n}{\hat{\epsilon}_r(n) \sqrt{K(n)} \log K(n)} . \quad (26)$$

Proof: We can use the previous lemma with our result on list codes to achieve the desired tradeoff. Let $\rho(n) = C^{\text{dep}}(\Lambda) - R_L$, where R_L is the rate for a list code of list size L . For the randomized code construction in Lemma 3 and the bound on the rate loss from Theorem 1, the error

probability is dominated by

$$\delta' \leq \frac{2LnR}{\sqrt{K} \log K} < \frac{8nC^{\text{dep}}(\Lambda) \log |\mathcal{Y}|}{\rho(n)\sqrt{K} \log K}. \quad (27)$$

Thus for large enough n we can bound the error $\hat{\varepsilon}_r(n)$ by $2\delta'$ and:

$$\rho(n) \leq \frac{16nC^{\text{dep}}(\Lambda) \log |\mathcal{Y}|}{\rho(n)\sqrt{K} \log K}. \quad (28)$$

■

This theorem gives some tradeoffs between error decay, key size, and rate loss. We could equally well phrase the result by fixing $K(n)$ first and finding the corresponding expressions.

V. DISCUSSION

In this paper we investigated arbitrarily varying channels in which the state sequence may be chosen based on the transmitted codeword. The key benefit derived from randomized coding is that the *decoding region* $D_{i,k}$ is still unknown to the jammer. This allows the encoder still achieve rates close to $C^{\text{dep}}(\Lambda)$, which is the natural upper bound.

The simplest examples of additive cost-constrained AVCs are the binary additive channel and the binary erasure channel. For both of these cases, the list coding problem for adversarial noise has been studied by many authors [4], [5], [15]–[17]. The results given in this paper above apply to more general channels with discrete alphabets and constrained adversaries.

As an example, let $\mathcal{X} = \{-1, 1\}$ and let $\mathcal{S} = \{-a, -a + 1, \dots, a\}$ for some integer a . The output Y of this channel is given by

$$Y = X + S. \quad (29)$$

That is, Y is the real addition of the input and state. We consider a power constraint on the jammer:

$$l(s) = s^2. \quad (30)$$

This is similar to the model studied by Shamai and Verdú [18], which analyzes a game between power constrained noise and an encoder with binary inputs. The results here can show how randomized coding can help even when the transmitted codeword is known to the jammer.

Consider the case when $K(n) = n^2$. In this case we have a rate loss of

$$16C^{\text{dep}}(\Lambda) \log(2a + 3) \cdot \frac{1}{2\hat{\varepsilon}_r(n) \log n} - \frac{1 \log n}{2n}. \quad (31)$$

So we can choose $\hat{\varepsilon}_r(n) = (\log n)^{-1/2}$ and have the rate loss and error probability converge to 0, although the latter will decay quite slowly. This is contrast to the case where jammer does not know the codebook, in which an error scaling of $1/n$ is achievable.

The problem studied in this paper was motivated by the problem of *online adversaries* who view the codeword causally and can attempt to modify their jamming situation accordingly. For these situations, the error performance using randomized codes with small keys should lie between the two

points above. One interesting question is how partial codeword knowledge can affect the error.

Another interesting extension is to the Gaussian scenario with a power constraint on the jammer and transmitter power constraint Γ . In [10] it was shown that key sizes with $K(n)/n \rightarrow \infty$ were sufficient to achieve the *randomized coding capacity*

$$C_r(\Lambda) = \frac{1}{2} \log \left(1 + \frac{\Gamma}{\Lambda} \right). \quad (32)$$

If the codeword is known to the jammer, then the results of [6] imply that randomized coding can only achieve

$$\frac{1}{2} \log \left(\frac{\Gamma}{\Lambda} \right). \quad (33)$$

The ‘‘Gaussian version’’ of our results here would give the same capacity value but with limited common randomness. It would be interesting to see if a similar list coding construction could be used in that setting as well.

ACKNOWLEDGMENT

This work was supported in part by the National Science Foundation under award CCF-0347298.

REFERENCES

- [1] D. Blackwell, L. Breiman, and A. Thomasian, ‘‘The capacities of certain channel classes under random coding,’’ *Annals of Mathematical Statistics*, vol. 31, pp. 558–567, 1960.
- [2] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Budapest: Akadémiai Kiadó, 1982.
- [3] R. Ahlswede and J. Wolfowitz, ‘‘Correlated decoding for channels with arbitrarily varying channel probability functions,’’ *Information and Control*, vol. 14, pp. 457–473, 1969.
- [4] M. Langberg, ‘‘Private codes or succinct random codes that are (almost) perfect,’’ in *Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2004)*, Rome, Italy, 2004.
- [5] A. Smith, ‘‘Scrambling adversarial errors using few random bits, optimal information reconciliation, and better private codes,’’ in *Proceedings of the 2007 ACM-SIAM Symposium on Discrete Algorithms (SODA 2007)*, 2007.
- [6] M. Agarwal, A. Sahai, and S. Mitter, ‘‘Coding into a source: a direct inverse rate-distortion theorem,’’ in *45th Annual Allerton Conference on Communication, Control and Computation*, 2006.
- [7] I. Csiszár and P. Narayan, ‘‘Arbitrarily varying channels with constrained inputs and states,’’ *IEEE Transactions on Information Theory*, vol. 34, no. 1, pp. 27–34, 1988.
- [8] —, ‘‘The capacity of the arbitrarily varying channel revisited : Positivity, constraints,’’ *IEEE Transactions on Information Theory*, vol. 34, no. 2, pp. 181–193, 1988.
- [9] R. Ahlswede, ‘‘Elimination of correlation in random codes for arbitrarily varying channels,’’ *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, vol. 44, no. 2, pp. 159–175, 1978.
- [10] A. Sarwate and M. Gastpar, ‘‘Randomization bounds on gaussian arbitrarily varying channels,’’ in *Proceedings of the 2006 International Symposium on Information Theory*, Seattle, WA, 2006.
- [11] R. Ahlswede, ‘‘Channels with arbitrarily varying channel probability functions in the presence of noiseless feedback,’’ *Zeitschrift für Wahrscheinlichkeit und verwandte Gebiete*, vol. 25, pp. 239–252, 1973.
- [12] —, ‘‘The maximal error capacity of arbitrarily varying channels for constant list sizes,’’ *IEEE Transactions on Information Theory*, vol. 39, no. 4, pp. 1416–1417, 1993.
- [13] A. Sarwate and M. Gastpar, ‘‘Deterministic list codes for state-constrained arbitrarily varying channels,’’ September 2007, submitted to the IEEE Transactions on Information Theory. [Online]. Available: <http://sipce.eecs.berkeley.edu/>
- [14] T. Cover and J. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

- [15] P. Elias, "List decoding for noisy channels," in *Wescon Convention Record, Part 2*. Institute of Radio Engineers (now IEEE), 1957, pp. 94–104.
- [16] V. Guruswami, J. Høastad, M. Sudan, and D. Zuckerman, "Combinatorial bounds for list decoding," *IEEE Transactions on Information Theory*, vol. 48, no. 5, pp. 1021–1034, 2002.
- [17] V. Guruswami, "List decoding from erasures: Bounds and code constructions," *IEEE Transactions on Information Theory*, vol. 49, no. 11, pp. 2826–2833, 2003.
- [18] S. Shamai (Shitz) and S. Verdú, "Worst-case power-constrained noise binary-input channels," *IEEE Transactions on Information Theory*, vol. 38, no. 5, pp. 1494–1511, 1992.