PROCEEDINGS OF THE 19TH ANNUAL CONFERENCE ON INFORMATION SCIENCES AND SYSTEMS, THE JOHNS HOPKINS UNIVERSITY, PAGES 722-725, MARCH 1985

TR 85-44

INTERLEAVING AND CHANNELS WITH UNKNOWN MEMORY †

Brian Hughes and Prakash Narayan

Department of Electrical Engineering
University of Maryland
College Park, Maryland 20742

ABSTRACT

Communications channels which have unknown, time-varying parameters frequently arise in practice. In certain instances, such as in jammed channels, the duration of memory is not known, and may also be time-varying. Most models of jammed channels which appear in the literature, however, assume that the jammer is restricted to *memoryless* jamming techniques; this restriction is usually rationalized by arguing that the use of **interleaving** by the transmitter is sufficient to 'destroy' channel memory. In this paper, the ability of interleaving to eliminate channel memory is investigated by comparing the capacity of jammed channels *with memory and random interleaving* with that of similar *memoryless*, jammed channels.

The **Compound Channel (CC)** and the **Arbitrarily Varying Channel (AVC)** are models of jammed channels ; the former admits only *fixed*, *memoryless* jamming techniques, while the latter also permits *time-varying* strategies with *memory*. Our models slightly generalize those in : we impose a cost structure on the collection of jamming techniques. Three types of cost constraints are distinguished: *peak symbol, peak codeword* and *average* constraints.

For each type of cost constraint, the performance which is achievable using determinstic codes with random interleaving is equal to the performance achievable using *random codes*. For cost constraints of the *peak* type (also, for no cost at all), the coding problem for the AVC with interleaving is shown to be *asymptotically equivalent* to the corresponding problem for the compound channel. The capacities of these two models are therefore the same. We conclude that, under peak cost constraints, interleaving does indeed asymptotically 'eliminate' channel memory.

In contrast, for the *average* cost constraint, it is shown that these capacities are *not* the same. We conclude that, in this case, interleaving does not (even in an asymptotic sense) result in a coding channel which is memoryless. Further, the performance of a given code on the AVC with interleaving is *always worse (and typically is much worse)* than the performance of the same code on the CC.

1. Introduction

Channels with an incomplete statistical description have long been of interest in the study of jammed communications. In the information-theoretic literature, two models predominate: the **compound channel (CC),** and the **arbitrarily varying channel (AVC)** (cf. Csissar and Korner [3]). The former admits only *stationary, memoryless* jamming signals, while the latter also permits *time-varying* signals with *memory*. The compound channel was independently introduced by Blackwell, Brieman and Thomasian [4] , Dobrushin [5] , and Wolfowitz [6] ; all of whom determined the capacity. AVCs were introduced by Blackwell, Breiman and Thomasian [7] , who established the capacity over the class of *random codes*. One interesting feature of AVCs is that random coding, deterministic coding with a maximal probability of error concept, and deterministic coding with an average probability of error concept all lead to different capacities (in general). Although much is now known about the latter two problems (see [3] , chapter 6), general computable characterizations of these capacities remain unknown.

Although, in most cases, the AVC offers the more realistic model of jamming, the CC has been used almost exclusively to study applications (e.g. Viterbi and Jacobs [8] , McEliece [9] , and McEliece and Stark[10]). The intuitive justification which is usually given for this choice of modeling is based on the use of *interleaving* in the coding system. An *interleaver* is a device which permutes the temporal order of code symbols prior to transmission; a *deinterleaver* applies the inverse permutation to the received symbols to restore the original order. It is argued that appending an interleaver before, and a deinterleaver after, a jammed channel results in a combined system which is well-modeled by the CC model.

In this paper, we study coding for the AVC with random interleaving. Our CC and AVC models slightly generalize those of the above authors: we impose a cost structure on the set of jamming symbols. In particular, three types of jamming cost constraints are considered: *peak symbol, peak codeword* and *average*. For each type of constraint, the λ-capacity (or capacity, should it exist) is characterized. In addition, we determine under what conditions, and in what sense, this channel is approximated by a CC model. Our main results can be summarized as follows. For each type of cost constraint, the performance which is achievable using deterministic codes with random interleaving on the AVC is equal to the performance achievable using random codes. For cost constraints of the peak type (symbol or codeword) we can further equate the coding problem to that of a corresponding compound channel. In contrast, for the average cost constraint it is shown that there is no equivalent compound channel problem.

† The research of the first author was sponsored by the Office of Naval Research under grant no. N00014-84-6-0101; the research of the second author was supported by the National Science Foundation under grant no. NSF ECS-82-0444-9 and by the Minta Martin Fund for Aerospace Research from the University of Maryland.

2. Definitions and a Summary of Results

Throughout this paper, we consider the following channel:

$$\left\{ w(y \mid x \mid s) \,\middle|\, y \in A_y,\, x \in A_x,\, s \in A_s \right\}, \tag{2.1}$$

where A_x and A_y (but not necessarily A_s) are finite sets and

$$w(y \mid x \mid s) \geq 0 \quad \text{for all } y \in A_y,\, x \in A_x,\, s \in A_s\,,$$
$$\sum_{y \in A_y} w(y \mid x \mid s) = 1 \quad \text{for all } x \in A_x,\, s \in A_s\,.$$

Let $\mathbf{R}$ denote the real line and Σ_R the Borel sets on $\mathbf{R}$. Let Σ_s be a σ-algebra of subsets of A_s so that $w(y \mid x \mid \cdot)$ is Σ_s–Σ_R measurable for each $y \in A_y$ and $x \in A_x$. For each $\mathbf{y}=(y_1, \ldots, y_n) \in A_y^n$, $\mathbf{x}=(x_1, \ldots, x_n) \in A_x^n$, $\mathbf{s}=(s_1, \ldots, s_n) \in A_s^n$, we define

$$w^n(\mathbf{y} \mid \mathbf{x} \mid \mathbf{s}) = \prod_{i=1}^{n} w(y_i \mid x_i \mid s_i)\,. \tag{2.2}$$

Let Σ_s^n be the σ-algebra on A_s^n generated by the cylinder sets with bases in Σ_s. For $\mathbf{S}_n$, an A_s^n-valued random variable measurable with respect to Σ_s^n, define

$$w_n(\mathbf{y} \mid \mathbf{x} \mid \mathbf{S}_n) = \mathbf{E}\, w^n(\mathbf{y} \mid \mathbf{x} \mid \mathbf{S}_n)\,, \tag{2.3}$$

where $\mathbf{E}$ denotes mathematical expectation. For any random variable S measurable on (A_s, Σ_s), define $r_n(S) = (S_1, \ldots, S_n)$ to be the A_s^n-valued random variable all of whose components are independent replications of S.

We now define several types of coding systems which will be studied in this paper. By an $(n\,,M)$ *deterministic (block) code*, C_n^D, we mean a system

$$\{(\mathbf{u}_1, D_1), \ldots, (\mathbf{u}_M, D_M)\}\,, \tag{2.4}$$

where $\mathbf{u}_i \in A_x^n$ for each i, and $\{D_i\}_{i=1}^{M}$ are disjoint subsets of A_y^n. By an $(n\,,M)$ *random code*, C_n^R, we mean a random variable which takes values in the set of all $(n\,,M)$ codes; we denote this by

$$\{(\mathbf{u}_1^*, D_1^*), \ldots, (\mathbf{u}_M^*, D_M^*)\}\,. \tag{2.5}$$

An $(n\,,m\,,M)$ *interleaved code*, which we denote by C_n^I, consists of a pair $\{C_n^D, \pi\}$, where C_n^D is an $(n\,,M)$ deterministic code, and $\pi = (\pi_1, \ldots, \pi_{mn})$ is a random variable which specifies how the order of the symbols of m successive codewords are to be permuted. The random variable π takes values on the set of all distinct permutations of the sequence $(1, \ldots, mn)$. This coding system is used in the following way. The transmitter encodes m successive messages, producing a total of mn symbols, say $(x_1, \ldots, x_{mn})$. These are passed to the interleaver which performs the random experiment, π, and emits $(x_{\pi_1}, \ldots, x_{\pi_{mn}})$. The deinterleaver, which also has access to the experiment π, receives the output $(y_{\pi_1}, \ldots, y_{\pi_{mn}})$, and, before passing this data on to the decoder, restores the original ordering, $(y_1, \ldots, y_{mn})$, where y_i is the output corresponding to x_i. We ignore the obvious delays incurred by accumulating m codewords in the interleaver and the deinterleaver. Throughout this paper, we shall refer to the channel formed by grouping the interleaver, the channel $w(\cdot \mid \cdot \mid \cdot)$, and the deinterleaver together as the *coding channel.*

Having specified admissible coding systems, we now consider admissible jamming sequences. The AVC model admits any random jamming sequence, $\mathbf{S}_n = (S_1, \ldots, S_n)$, which is measurable on (A_s^n, Σ_s^n). The CC model admits only *stationary, memoryless* jamming sequences of the form $\mathbf{S}_n = r_n(S)$, for some (A_s, Σ_s) measurable S. In this paper, we generalize the usual AVC and CC models by further imposing a *cost structure* on the set of jamming symbols. In practice, some form of cost constraint - typically a power constraint - is usually present. Let $\phi(\cdot)$ be a continuous, non-negative, Σ_s–Σ_R measurable *cost function* on A_s, such that $\phi(s_0) = 0$ for some $s_0 \in A_s$. We shall consider three types of constraints on a sequence $\mathbf{S}_n = (S_1, \ldots, S_n)$:

(1) *peak symbol (PS):* $\phi(S_i) \leq c$ with probability one (w.p.1) for each i,

(2) *peak codeword (PC):* $\sum_{i=1}^{n} \phi(S_i) \leq nc$ w.p.1 ,

(3) *average (A):* $\mathbf{E}\phi(S_i) \leq c$ w.p.1 for each i,

where $c \geq 0$. Denote the set of $\mathbf{S}_n$ which satisfy PS, PC and A, respectively, by S_n^{PS}, S_n^{PC}, and S_n^{A}.

In this paper, a *coding problem* means a rule defining, for each n, *admissible codes*, together with a measure of the reliability of these codes. Since we shall consider many coding problems, it will be convenient to adopt a simple ternary nomenclature to distinguish them, viz.

$$\alpha \mid \beta \mid \gamma\,,$$

where α designates the type of coding system used: deterministic (D), random (R), or interleaved (I); β defines the channel model being used: the compound channel (CC) or the arbitrarily varying channel (AVC); and γ specifies the type of cost constraint imposed on the jammer: peak symbol (PS), peak codeword (PC) or average (A).

We are concerned primarily with the following cases:

- D|CC|PS, D|CC|A
- R|AVC|PS, R|AVC|PC, R|AVC|A
- I|AVC|PS, I|AVC|PC, I|AVC|A.

Remark: The constraint PC is not essentially different from PS for compound channel problems; so the case D|CC|PC is omitted. Cases of the form R|CC|γ are known to be equivalent to the corresponding cases D|CC|γ; this is a fortiori true of cases I|CC|γ as well. The problems D|AVC|γ are of interest, but are beyond the scope of the present paper. D|AVC|PS is equivalent to the usual deterministic coding problem for the AVC.

We now define the reliability measures to be used for each coding problem enumerated above. For cases D|CC|γ, define

$$\lambda_{CC}^{\gamma}(C_n^D) = \max_{1 \leq i \leq M} \; \sup_{S \in S_1^{\gamma}} \; w_n(\bar{D}_i \mid \mathbf{u}_i \mid r_n(S))\,, \tag{2.6}$$

where $\gamma =$ PS or A. For cases of the form R|AVC|γ, we define

$$\lambda_{AVC}^{\gamma}(C_n^R) = \max_{1 \leq i \leq M} \; \sup_{\mathbf{S}_n \in S_n^{\gamma}} \; \mathbf{E}\, w^n(\bar{D}_i^* \mid \mathbf{u}_i^* \mid \mathbf{S}_n)\,. \tag{2.7}$$

Finally, for the problems I|AVC|γ, we define

$$\lambda_{AVC}^{\gamma}(C_n^I) = \max_{1 \leq i \leq M} \; \sup_{\mathbf{S}_{mn} \in S_{mn}^{\gamma}} \; w^n(\bar{D}_i \mid \mathbf{u}_i \mid \pi_n(\mathbf{S}_{mn})) \tag{2.8}$$

where γ = PS, PC or A and where

$$\pi_n(\mathbf{s}_{mn}) = (S_{\pi_1^{-1}}, \ldots, S_{\pi_n^{-1}}) . \qquad (2.9)$$

Note that the jamming constraint in (2.8) is enforced over the block of mn interleaved symbols, rather than over the n code symbols as in (2.7). This, of course, makes no difference for constraints PS and A; however, for PC it is slightly weaker (i.e. admits more jamming sequences).

We say that a given code, C_n^α, is an (n, M, λ) code for the problem $\alpha \mid \beta \mid \gamma$, if the code is of the type α, and

$$\lambda_\beta^\gamma(C_n^\alpha) \leq \lambda .$$

A pair, (R, λ), where $R \geq 0$ and $0 \leq \lambda < 1$, is said to be *achievable* $\alpha \mid \beta \mid \gamma$, if for all $\epsilon > 0$ there exist, for all sufficiently large n (and m if α = I), an (n, M, λ') code for the problem $\alpha \mid \beta \mid \gamma$ such that

$$\log_2 M \geq n(R - \epsilon) \qquad (2.10a)$$

and

$$\lambda' \leq \lambda + \epsilon . \qquad (2.10b)$$

Let $C_{\alpha|\beta|\gamma}(\lambda)$, for $0 \leq \lambda < 1$, denote the largest R for which (R, λ) is achievable $\alpha \mid \beta \mid \gamma$; we call this the λ-*capacity* (see [2], section 7.7) of the problem. If, as is usually the case, $C_{\alpha|\beta|\gamma}(\lambda) = C_{\alpha|\beta|\gamma}$, a constant for all $0 \leq \lambda < 1$, then we call this the *capacity* of the problem.

As stated earlier, this paper is concerned primarily with the coding problem for AVCs with random interleaving, and its relationship to the random coding problem for AVCs and the deterministic coding problem for CCs. We now summarise our results. To simplify their statement, we introduce the following notation. Let X and S be independent random variables taking values in the sets A_x and A_s, respectively. Define $Y(S)$ to be the A_y-valued random variable which is related to X and S as follows:

$$Pr\{Y(S)=y \mid X=x, S=s\} = w(y \mid x \mid s) .$$

Define the *mutual information* in the usual way:

$$I(X, Y(S)) = \sum_{\substack{x \in A_x \\ y \in A_y}} Pr\{X=x, Y(S)=y\} \log_2 \left\{ \frac{Pr\{Y(S)=y \mid X=x\}}{Pr\{Y(S)=y\}} \right\} .$$

For the purposes of comparison, let us first present two theorems which characterize the capacities of those problems of the form D|CC|γ and R|AVC|γ.

Theorem 1: (Blackwell-Breiman-Thomasian-Dobrushin-Wolfowitz) For the cases D|CC|PS and D|CC|A capacities exist; these are given by

$$C_{D|CC|PS} = \max_X \inf_{S:\phi(S) \leq c} I(X, Y(S)) \qquad (2.11)$$

and

$$C_{D|CC|A} = \max_X \inf_{S:\mathbf{E}\phi(S) \leq c} I(X, Y(S)) . \qquad (2.12)$$

Remark: Although the cost structure used here is absent from the work of these authors [4], [5], [6], for compound channels the distinction is only notational.

Theorem 2: The coding problems R|AVC|PS and R|AVC|PC possess capacities which are given by

$$C_{R|AVC|PS} = \max_X \inf_{S:\phi(S) \leq c} I(X, Y(S)) \qquad (2.13)$$

and

$$C_{R|AVC|PC} = \max_X \inf_{S:\mathbf{E}\phi(S) \leq c} I(X, Y(S)) . \qquad (2.14)$$

In the case R|AVC|A, only a λ-capacity exists which is given by

$$C_{R|AVC|A}(\lambda) = \max_X \inf_{S:\lambda\mathbf{E}\phi(S) \leq c} I(X, Y(S)) . \qquad (2.15)$$

Remark: (2.13) is due to Blackwell, Breiman and Thomasian [7]; (2.14) and (2.15) are new.

Let us now consider the coding problems which involve interleaved codes. Suppose that a deterministic code with random interleaving is used on the channel (2.1). When the jammer emits a given sequence, say $\mathbf{s}_{mn}$, the coding channel appears to be jammed by the sequence $\pi_n(\mathbf{s}_{mn})$, defined in (2.8). Of course, $\pi_n(\mathbf{s}_{mn})$ is not in general memoryless; however, the next theorem shows that, for an appropriate choice of π, the random sequences $\pi_n(\mathbf{s}_{mn})$ and $r_n(\pi_1(\mathbf{s}_{mn}))$ lead to essentially the same error probabilities, for large m.

Theorem 3: Consider the collection of channels in (2.1); let $\mathbf{s}_{mn}$ be any sequence in A_s^{mn}. Let $\hat{\pi} = (\hat{\pi}_1, \ldots, \hat{\pi}_{mn})$ be a random variable which is uniformly distributed on the set of all distinct permutations of the sequence $(1, \ldots, mn)$. Suppose that for some $\mathbf{u} \in A_x^n$ and some $D \subset A_y^n$ one has

$$w^n(D \mid \mathbf{u} \mid r_n(\hat{\pi}_1(\mathbf{s}^{mn}))) \geq \delta > 0 , \qquad (2.16)$$

then

$$\left| \frac{w_n(D \mid \mathbf{u} \mid \hat{\pi}_n(\mathbf{s}^{mn}))}{w_n(D \mid \mathbf{u} \mid r_n(\hat{\pi}_1(\mathbf{s}^{mn})))} - 1 \right| < \epsilon(\delta, n, m) , \qquad (2.17)$$

where the right-hand expression depends only on $|A_x|$ and $|A_y|$ and not on A_s or $w(\cdot \mid \cdot \mid \cdot)$. Further, $\epsilon(\delta, n, m) \to 0$ as $m \to +\infty$, and $\epsilon(\delta, n, m_n) \to 0$ as $n \to +\infty$ provided $m_n \geq (n \ln n)^2 (n^2+1)^{|A_x||A_y|}$ for all large n.

Remark: The lemma implies that, if D is a good decoding set for $\mathbf{u}$ when the memoryless version of $\mathbf{s}^{mn}$ is used, namely $r_n(\hat{\pi}_1(\mathbf{s}^{mn}))$, then it will be good for $\hat{\pi}_n(\mathbf{s}^{mn})$ as well, provided m is sufficiently large compared to n.

Remark: Some complexity issues arise here which we make little attempt to address. First, it is likely that the bound on m_n can be made smaller and the theorem will still hold. Our interest, however, lies in the coding theorems below, for which these estimates suffice. Second, the interleaver, $\hat{\pi}$, has $(mn)!$ equiprobable realisations, which becomes unmanageable even for moderate m and n. Using an argument similar to that used in Ahlswede [11], however, it can easily be shown that a random interleaving strategy having only n^2 realizations will suffice to yield the lemma (with perhaps a different bound on m_n).

Although interleaved codes are a form of random coding, they are structurally much more restricted than general random codes. Since it is known (cf. [3] that deterministic codes lead to a lower capacity than random codes for AVCs, it would be reasonable to expect a similar result for interleaved codes. However as the following theorem shows, no such reduction in achievable rates occurs.

Theorem 4: The λ-capacity (or capacity, should it exist) of the coding problem I|AVC|γ equals the λ-capacity (or capacity) of the coding problem R|AVC|γ.

The significance of Theorem 4 is that any performance which can be achieved using random codes on the AVC can also be achieved by a coding sytem which consists of a deterministic code in series with a random interleaver. In our view, this result greatly enhances the practical value of random coding theorems for AVCs. Although a general random code might be very difficult to implement, this is not true of interleaved codes. In practice, a random interleaver might be realized by a *pseudorandom interleaver , for which a hardware implentation already exists.*

We now address the question of whether the coding channel which results when interleaving is used on an AVC approximates (in some sense) a compound channel. From Theorem 1, 2, and 4 it is clear that problems D|CC|A and I|AVC|PS have identical capacities; the same is true of problems D|CC|A and I|AVC|PC. However, using Theorem 3, we can prove the following much stronger result.

Theorem 5: Let C_n^D be *any* deterministic code for the channel (2.1), and let an interleaved code, C_n^I, be formed by combining C_n^D with the interleaver π of Theorem 3. Then there exists a function, $\delta(m, n, |A_x|, |A_y|)$ which does not depend on $w(\cdot|\cdot|\cdot)$ or A_x or the code C_n^D, so that

$$\lambda_{CC}^{PS}(C_n^D) \leq \lambda_{AVC}^{PS}(C_n^I) \leq \lambda_{CC}^{PS}(C_n^D) + \delta(m, n, |A_x|, |A_y|),$$

and

$$\lambda_{CC}^{A}(C_n^D) \leq \lambda_{AVC}^{PC}(C_n^I) \leq \lambda_{CC}^{A}(C_n^D) + \delta(m, n, |A_x|, |A_y|),$$

where $\delta(m, n, |A_x|, |A_y|) \to 0$ as $m \to +\infty$.

Thus, as m increases, the error probability incurred by any code on the AVC with constraint PS (respectively, PC) when interleaving is used, is asymptotically equal to the error probability incurred on the compound channel with constraint PS (resp. A). Thus, the coding channel is, for large m, essentially a compound channel.

It only remains to examine the case I|AVC|A, which, according to Theorem 4, has the λ-capacity given in (2.15). This obviously corresponds to no compound channel we have considered; in fact, we can further assert that it cannot correspond to any other compound channel since the latter must have a strict capacity (cf. Wolfowitz [2] , section 7.7). Although the interleaver asymptotically renders the coding channel ***stationary***, it does not render it ***memoryless***. It is of particular interest to compare the cases D|CC|A and I|AVC|A, since, in practice, it is common for D|CC|A to be used as an approximation to I|AVC|A. Note that for all $\lambda < 1$ one has

$$C_{I|AVC|A}(\lambda) \leq C_{D|CC|A}$$

In many practical cases $C_{I|AVC|A}(0+) = 0$, so that if the case I|AVC|A were to be analyzed using D|CC|A as a model, one would mistakenly conclude that for any $R < C_{D|CC|A}$ arbitrarily reliable codes exist; whereas in fact the error probability is bounded away from zero.

8. Acknowledgements

We wish to thank Anthony Ephremides for many stimulating discussions on this problem.

References

1. G. C. Clark, Jr. and J. B. Cain, *Error-Correction Coding for Digital Communications*, Plenum, New York, 1981.
2. J. Wolfowitz, *Coding Theorems of Information Theory, 3rd ed.*, Springer-Verlag, Berlin, 1978.
3. I. Csiszar and J. Korner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*, Academic Press, New York, 1981.
4. D. Blackwell, L. Breiman, and A. J. Thomasian, "The Capacity of a Class of Channels," *Ann. Math. Stat.*, vol. 30, pp. 1229-1241, 1959.
5. R. L. Dobrushin, "Optimal Information Transmission over a Channel with Unknown Parameters," *Radiotekh. Elektron.*, vol. 4, no. 12, pp. 1951-1956, 1959.
6. J. Wolfowitz, "Simultaneous Channels," *Arch. Ration. Mech. Anal.*, vol. 4, no. 4, pp. 371-386, 1960.
7. D. Blackwell, L. Breiman, and A. J. Thomasian, "The Capacities of Certain Channel Classes under Random Coding," *Ann. Math. Stat.*, vol. 31, pp. 558-567, 1960.
8. A. J. Viterbi and I. M. Jacobs, "Advances in Coding and Modulation for Noncoherent Channels Affected by Fading, Partial-Band, and Multiple-Access Interference," *Advances in Communications Systems*, vol. 4, pp. 279-308, Academic Press, New York, 1975.
9. R. J. McEliece, "Communications in the Presence of Jamming - An Information-Theoretic Approach," in *Secure Digital Communications*, ed. G. Longo, pp. 127-166, Springer-Verlag, New York, 1983.
10. R. J. McEliece and W. E. Stark, "An Information-Theoretic Study of Communication in the Presence of Jamming," *Proc. IEEE International Conference on Communications*, pp. 45.3.1-45.3.5, 1981.
11. R. Ahlswede, "Elimination of Correlation in Random Codes for Arbitrarily Varying Channels," *Z. Wahrscheinlichkeitstheorie Verw. Gebiete*, vol. 44, pp. 159-175, 1978.