An Information-Theoretic and Game-Theoretic Study of Timing Channels

James Giles, *Member, IEEE,* and Bruce Hajek, *Fellow, IEEE*

***Abstract*—This paper focuses on jammed timing channels. Pure delay jammers with a maximum delay constraint, an average delay constraint, or a maximum buffer size constraint are explored, for continuous-time or discrete-time packet waveforms. Fluid waveform approximations of each of these classes of waveforms are employed to aid in analysis. Channel capacity is defined and an information-theoretic game based on mutual information rate is studied. Min-max optimal jammers and max-min optimal input processes are sought. Bounds on the min–max and max–min mutual information rates are described, and numerical examples are given. For maximum-delay-constrained (MDC) jammers with continuous-time packet waveforms, saddle-point input and jammer strategies are identified. The capacity of the maximum-delay constrained jamming channel with continuous-time packet waveforms is shown to equal the mutual information rate of the saddle point. For MDC jammers with discrete-time packet waveforms, saddle-point strategies are shown to exist. Jammers which have quantized batch departures at regular intervals are shown to perform well. Input processes with batches at regular intervals perform well for MDC or maximum-buffer-size-constrained jammers.**

***Index Terms*—Channel coding, covert timing channels, jamming, network security.**

I. Introduction

PACKET timing channels can sometimes be used to convey covert messages. For example, as shown in Fig. 1, Bob can transmit packets containing innocuous messages to Alice while transmitting secret messages using packet timing. While the channel may be naturally noisy due to delays in the underlying packet channel; it still might be possible for Bob and Alice to communicate at a high data rate using timing. Even if the communication is monitored, a secret message may be overlooked.

We define a covert channel to be any channel used for communication that is either not intended to be used for communication or that is intended to be used in a fundamentally different way. A *covert timing channel* is a covert channel that uses timing to convey information. See [32] for a collection of other definitions.

Manuscript received September 11, 2000; revised May 22, 2002. This work was supported by the Department of Defense under NDSEG Graduate Fellowship and by the National Science Foundation under Contract NSF ANR 99-80594.

J. Giles was with the Coordinated Science Laboratory and the Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign. He is now with IBM T. J. Watson Research Center, Hawthorne, NY 10532 USA (e-mail: jamesgiles@watson.ibm.com).

B. Hajek is with the Coordinated Science Laboratory and the Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA (e-mail: b-hajek@uiuc.edu).

Communicated by P. Narayan, Associate Editor for Shannon Theory.

Publisher Item Identifier 10.1109/TIT.2002.801405.

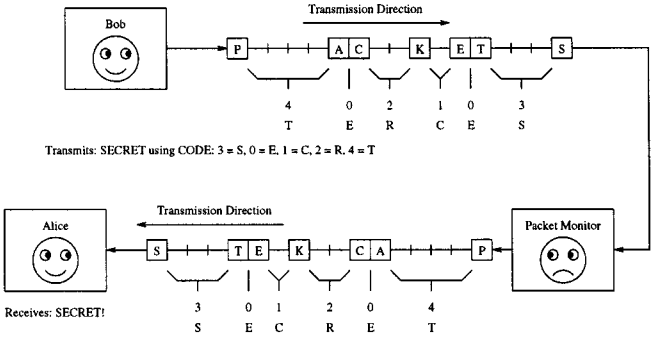

Fig. 1. Motivating example.

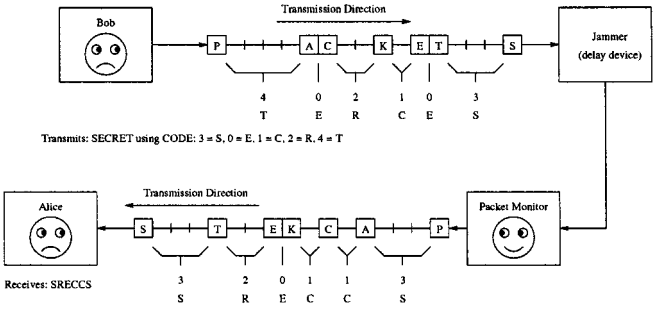

Fig. 2. Motivating example.

Covert timing channels can exist due to processor sharing on a single machine, shared message buffers, or standard network connections. In a typical covert channel scenario, a trusted user with access to sensitive data passes the data to a user who does not have access to the sensitive data by using a shared resource as a covert channel. The trusted user may be unaware that data is being passed if he is running a program containing a Trojan horse.

Taking the viewpoint of a security authority who expects to be able to monitor all communications on a given packet communications system, the existence of this covert packet timing channel may be undesirable. In particular, the complexity of monitoring all possible ways in which information can be conveyed through packet timing may make monitoring impractical. Rather than trying to monitor all packet timing channels, one solution is to employ jamming devices which limit the usefulness of the packet timing channels, as shown in Fig. 2. If the rate at which information can be conveyed using a packet timing channel in the presence of a jamming device is guaranteed to be acceptably low, the security authority will not need to monitor the packet timing channel.

On the other hand, taking the viewpoint of a network user concerned with secrecy, the use of a packet timing channel to

circumvent packet monitoring systems may be desirable. Of course, encryption is a straightforward way to attempt privacy, but encryption can draw attention and can be compromised with sufficient effort. The use of a packet timing channel, possibly with encryption, may be a way to communicate privately without drawing attention.

The goal of this work is to develop good jamming and coding strategies for timing channels and to find bounds on the information leakage under such strategies. We assume that all noise in the timing channel is introduced by an intelligent jammer, since the capacity of a timing channel with naturally occurring delay is at least as large as the capacity of a jammer timing channel assuming the same reasonable constraints on the jammer.

Perhaps Gallager [16] was the first to study the information that packets can convey beyond the information within the traditional data portion of the packets. There is a literature on covert communication through timing channels [3], [4], [17], [19]–[21], [26]–[31], [33], [35], [36], [38] and through storage channels [9], [33], [37]. Covert timing channels present a unique security problem in that there is no apparent way to completely eliminate them in a reliable communication system (e.g., [30]). As a method to combat covert timing channels, Hu [19] proposes to make all clocks available to user processes on computers noisy.

One computer can covertly communicate to another by modulating the timing of acknowledgment packets. Kang and Moskowitz [20] introduced a mechanism that reduces the capacity of this channel by smoothing and randomizing the delay of acknowledgment packets. Venkatraman and Wolfe [38] discuss estimation of the capacity of a covert channel using an adaptive scheduling policy and they discuss the auditability of network covert channels based on changes in traffic volume over time.

The work of Anantharam and Verdú [3] and Bedekar and Azizoglu [4] can be viewed as a study of queues as timing noise devices. Sundaresan and Verdú [35] found how to minimize the capacity of a timing channel consisting of a single server queue. They constrained the packet service times, rather than the total delay through the queue.

The remainder of the paper is organized as follows. Section II presents the channel models and assumptions, and introduces two generic jamming channels and two generic input processes. Section III presents relationships among the various channel models and gives a capacity bound. Sections IV–VI consider the strategies applied to timing channels in the presence of jammers with a maximum delay constraint, jammers with a maximum buffer size constraint, and jammers with an average delay constraint, respectively. Discussion, examples, conclusions, and ideas for future research are provided in Section VII. Proofs of two theorems and additional information about jammers with an average delay constraint are given in the Appendix.

II. Problem Formulation

Waveforms representing cumulative arrivals are used in this paper to describe the input or output of a covert packet timing channel. A *continuous-time packet waveform* A is a right-continuous, nondecreasing, integer-valued function on the nonnegative real line with $A(0) = 0$. For $t \geq 0$, $A(t)$ represents the number of packets that arrive in the interval $[0, t]$. Multiple packets can arrive simultaneously for the continuous-time waveforms.

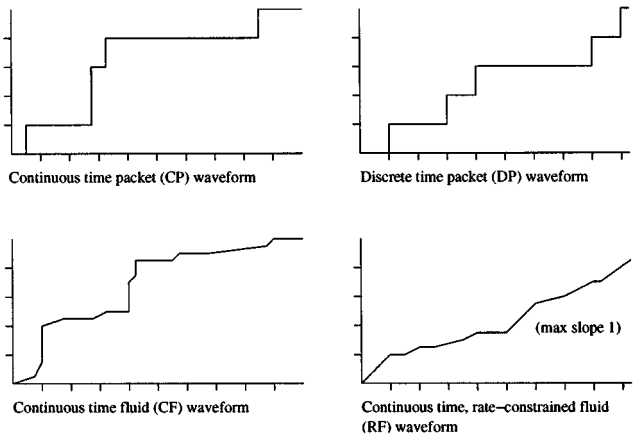

Fig. 3. Typical waveforms for various waveform classes.

A *discrete-time packet waveform* A is a function on the nonnegative integers such that $A(0) = 0$ and $A(k) - A(k-1) \in \{0, 1\}$ for all $k \geq 1$. Throughput this paper, when discrete-time models are concerned, for an integer $k \geq 1$, slot k means the same as time k. For integer $k \geq 1$, $A(k)$ is the number of packets that arrive in slots 1 through k, and $A(k) - A(k-1)$ is the number of packets, either zero or one, that arrive in slot k. The restriction of at most one packet arrival per slot introduces a lower bound on inter-packet spacing.

Two more classes of waveforms are defined to aid in analysis. A *continuous-time fluid waveform* A is a right-continuous, nondecreasing function on the nonnegative real line with $A(0) = 0$. Such waveforms can be obtained from the original continuous-time packet waveforms by letting the packet size converge to zero. Similarly, a *continuous-time, rate-constrained fluid waveform* A is a continuous-time fluid waveform such that $A(t) - A(s) \leq t - s$ for $0 \leq s < t$. These waveforms are obtained from the original discrete-time packet waveforms by letting the packet size and slot size converge to zero together. Given a fluid waveform A, and $a \geq 0$, let $T_A(a) = \inf\{t\colon A(t) \geq a\}$, so that $T_A(a)$ is the time by which a units of fluid arrive for waveform A.

Typical input waveforms for the various waveform classes are provided in Fig. 3. Note that the waveforms do not include packet sequence numbers. This reflects our focus on packet timing rather than on bits within the packets. In particular, we do not permit the transmitter to code covert information by using out-of-order sequence numbers, so the jammer has no incentive to reorder packets. Waveforms of a finite duration T are frequently considered in the paper. Those are either continuous-time waveforms defined over the interval of time $[0, T]$, or discrete-time waveforms defined over a finite interval in discrete time $\{0, 1, \ldots, T\}$.

Ideally, we would like to say something about actual codes and jamming strategies. That is, for a particular class of jamming strategies, we would like to find codes of high rate, with arbitrarily small probability of decoding error when passed through a jammer. In addition, we would like to understand the

most effective jamming strategies in a particular class. This view of the problem is called the *coding framework*.

An (M, T) *code* consists of an indexed set of M waveforms of duration T, $\{X_1, \ldots, X_M\}$, and a mapping ϕ from the set of duration T waveforms to $\{0, 1, \ldots, M\}$. The set $\{X_1, \ldots, X_M\}$ is the set of codewords, ϕ is the decoder, and decoder output 0 denotes an erasure. The rate of such a code is $\frac{1}{T} \log M$.

Let $\mathcal{W}$ represent a class of jammers and let $\lambda > 0$. A rate R is *achievable* by randomized codes for arrival rate λ codewords and the jammer class $\mathcal{W}$ if for any $\epsilon > 0$ there exists a sequence of random (M, T) codes such that the following three properties hold:

i) $\frac{1}{T} \log M \geq R - \epsilon$ for all T;

ii) $P\left(\left|\frac{X_m(T)}{T} - \lambda\right| \leq \epsilon\right) = 1$, for all T and $1 \leq m \leq M$;

iii) the maximal probability of decoding error over all jammers in the class and over all codewords converges to zero as $T \to \infty$.

The *capacity* C for $\mathcal{W}$ and λ is the supremum of such rates R.

We assume that a jammer can choose any causal delay strategy, including strategies that change packet ordering, subject to constraints on the delay. The jamming strategy can be deterministic, or it can be random with memory. However, in the case of packet waveforms, the jammer cannot delete packets or insert duplicates or additional packets since this might impact the underlying packet communication system. Similarly, a jammer cannot alter a fluid waveform by inserting or deleting fluid. The jammer knows that the number of packets or amount of fluid in each codeword is approximately the same, but in most cases the jammer does not know λ or the codewords. In addition, in most cases we assume that the jammer knows time zero for the coder (i.e., the time when the code transmission begins). We use randomized code capacity, which implies that the coder and decoder have access to a source of common randomness, so that they can randomly select codes without the jammer's knowledge.

The constraints considered for jammers include a maximum-delay constraint, maximum-buffer-size constraint, and average-delay constraint. For packet waveforms, a jammer is called a *maximum-delay-constrained* (MDC) jammer with delay parameter D if it delays no packet by D or more time units. For fluid waveforms, a jammer is MDC with delay parameter D if for any input X and corresponding output Y, $T_Y(a) < T_X(a) + D$ for all $a \geq 0$. For packet waveforms, a *maximum-buffer-constrained* (MBC) jammer with parameter B is a jammer that holds no more than B packets at any time in the continuous-time case, or holds no more than B packets from one slot to the next slot in the discrete time case. For fluid waveforms, an MBC jammer with parameter B cannot hold more than B units of fluid at any time. For packet waveforms, given the length T of a code to be used, a jammer is called an *average-delay-constrained* (ADC) jammer with delay parameter D if whenever $X(T) > 0$

$$\frac{1}{X(T)} \sum_{i=1}^{X(T)} D_i \leq D$$

where $X(T)$ is the number of arrivals up to time T and D_i is the delay of the ith packet. Similarly, for fluid waveforms, given the length T of a code to be used, a jammer is ADC with delay parameter D if whenever $X(T) > 0$,

$$\frac{1}{X(T)} \int_{a=0}^{X(T)} (T_Y(a) - T_X(a))\, da \leq D$$

where Y is the output waveform. When the delay parameter is not explicitly stated, the parameter D is used for MDC and ADC jamming channels and the parameter B is used for MBC jamming channels.

The coder and the decoder are aware of the particular delay constraints placed on the jammer, but are not aware of the actual strategy used by the jammer. In the case of an maximum-delay constraint, for example, the coder and the decoder will know that there is a maximum-delay constraint for the jammer and also know the value of the maximum delay. The coder does not receive feedback from the decoder. In practice, if the delay constraints are unknown to the coder and the decoder, they can make conservative assumptions. For example, in the case of a maximum-delay constraint, the coder and the decoder can assume a particular maximum delay, and as long as the actual maximum delay is less than or equal to the assumed maximum delay, the communication will be reliable. Similarly, in the case of an average-delay constraint, the coder and decoder can assume the average delay for the packets sent within a codeword is bounded by some number D, and as long as the actual average is less than or equal to D, the communication should be reliable.

We have just described the coding framework. An alternative is to view the input as a random process passing through a jamming channel, and to consider the mutual information rate between the input and output. We call this point of view the *information-theoretic framework*. For $T > 0$ and processes X and Y, let $\overline{I}_T(X;Y) = \frac{1}{T} I_T(X;Y)$ where $I_T(X;Y)$ is the mutual information for X and Y up to time T. For an input process X and a jamming channel W, we write $\overline{I}_T(X, W) = \overline{I}_T(X; W_X)$ where W_X is the output of jammer W when X is the input.

We consider a zero sum game between the encoder and the jammer, where mutual information per unit time is the objective function. We first describe the information-theoretic game for MDC and MBC jamming channels. We then describe the game for ADC jamming channels, which requires slightly different encoder and jammer constraints for technical reasons.

Fix an arrival rate $\lambda > 0$. For a fixed $T > 0$, inputs for MDC and MBC jamming channels in the information-theoretic framework are constrained to be in the class $\mathcal{X}_T = \{X : E[X(T)] = \lambda T\}$ where $X(T)$ is the number of packets (or quantity of fluid) in the input up to time T. We call input processes in $\mathcal{X}_T$ *rate* λ *input processes* for MDC or MBC jamming channels.

For $T > 0$ and $\mathcal{W}$, a particular class of MDC or MBC jamming channels, let

$$\underline{V}_T = \sup_{X \in \mathcal{X}_T} \inf_{W \in \mathcal{W}} \overline{I}_T(X, W)$$

and

$$\overline{V}_T = \inf_{W \in \mathcal{W}} \sup_{X \in \mathcal{X}_T} \overline{I}_T(X, W).$$

An input X^* for which

$$\inf_{W\in\mathcal{W}} \overline{I}_T(X^*, W) = \underline{V}_T \tag{1}$$

is called an *optimal* input on the interval $[0, T]$ in the sense that no matter what jamming strategy is used, at least $\underline{V}_T$ units of mutual information will get through the jammer when input X^* is used and no other input can guarantee more information leakage than $\underline{V}_T$. Likewise, a jammer W^* for which

$$\sup_{X\in\mathcal{X}_T} \overline{I}_T(X, W^*) = \overline{V}_T \tag{2}$$

is called an *optimal* jammer on the interval $[0, T]$ in the sense that no matter what input is used, at most $\overline{V}_T$ mutual information will get through the jammer, and no other jammer can guarantee less information leakage than $\overline{V}_T$. Assuming intelligent opponents, the encoder will use strategy X^* (if it exists) and the jammer will use strategy W^* (if it exists). Even if these strategies do not exist, intelligent encoders and jammers can select strategies which come arbitrarily close to satisfying (1) and (2), respectively.

For MDC and MBC jamming channels, we also define

$$\underline{V} = \liminf_{T\to\infty} \underline{V}_T \tag{3}$$

and

$$\overline{V} = \limsup_{T\to\infty} \overline{V}_T. \tag{4}$$

Slightly abusing notation, we refer to $\underline{V}$ as the max-min information rate and we refer to $\overline{V}$ as the min-max information rate. We believe that the study of $\underline{V}$ and $\overline{V}$ is informative for understanding the capacity of jammed timing channels. In particular, for many jamming channels, we believe that limits exist in (3) and (4) and that $\overline{V} = \underline{V} = C$. We show this result for the MDC jamming channel for continuous-time packet waveforms, and for degenerate cases where $\overline{V} = \underline{V} = C = \infty$. Note that in general, $\overline{V}_T \geq \underline{V}_T$ and $\overline{V} \geq \underline{V}$.

An input strategy is a sequence of inputs $(X^{T_n}: n \geq 1)$ with $\lim_{n\to\infty} T_n = \infty$, such that $X^{T_n} \in \mathcal{X}_{T_n}$ for each n. An input strategy X is max-min optimal if

$$\lim_{n\to\infty} \inf_{W\in\mathcal{W}} \overline{I}_{T_n}\left(X^{T_n}, W\right) \geq \limsup_{n\to\infty} \inf_{W\in\mathcal{W}} \overline{I}_{T_n}\left(\tilde{X}^{T_n}, W\right)$$

for any other input strategy $(\tilde{X}^{T_n}: n \geq 1)$. Similarly, a jamming strategy is a sequence of jamming channels $(W^{T_n}: n > 0)$ with $\lim_{n\to\infty} T_n = \infty$ such that $W^{T_n} \in \mathcal{W}$. A jamming strategy is called min-max optimal if

$$\lim_{n\to\infty} \sup_{X\in\mathcal{X}_{T_n}} \overline{I}_{T_n}\left(X, W^{T_n}\right) \leq \liminf_{n\to\infty} \sup_{X\in\mathcal{X}_{T_n}} \overline{I}_{T_n}\left(X, \tilde{W}^{T_n}\right)$$

for any other jamming strategy $(\tilde{W}^{T_n}: T_n > 0)$. If $\underline{V} = \overline{V}$, then the common value $\underline{V}$ or $\overline{V}$ is the saddle-point information rate. Any pair consisting of a max-min optimal input strategy and a min-max optimal jamming strategy is called a saddle point.

Next, we define the information-theoretic game for ADC jamming channels. For fixed $T > 0$ and $\epsilon > 0$, an input for an ADC jamming channel is constrained to be in the class

$$\mathcal{X}_{T,\epsilon} = \{X: P(|X(T) - \lambda T| \leq \epsilon T) = 1\}$$

where $X(T)$ is the number of packets (or amount of fluid) in the input up to time T. We call inputs in $\mathcal{X}_{T,\epsilon}$ *rate* $\lambda \pm \epsilon$ *inputs*. For a given class of arrival processes and fixed $T > 0$, we define the class of ADC packet jamming channels as

$$\mathcal{W}_T = \left\{W: P\left(\frac{1}{X(T)} \sum_{i=1}^{X(T)} D_i \leq D\right) = 1 \ \forall \text{ inputs } X\right\}$$

where $X(T)$ is the number of packets in the input up to time T and D_i is the delay added to the ith packet by a jammer W. The delay constraint for ADC fluid jammers is similarly defined.

For $T > 0$, $\epsilon > 0$, and a particular class of ADC jamming channels $\mathcal{W}_T$, let

$$\underline{V}_{T,\epsilon} = \sup_{X\in\mathcal{X}_{T,\epsilon}} \inf_{W\in\mathcal{W}_T} \overline{I}_T(X, W)$$

and

$$\overline{V}_{T,\epsilon} = \inf_{W\in\mathcal{W}_T} \sup_{X\in\mathcal{X}_{T,\epsilon}} \overline{I}_T(X, W).$$

For ADC jamming channels, max-min and min-max information rates

$$\underline{V}7 = \lim_{\epsilon\to 0} \liminf_{T\to\infty} \underline{V}_{T,\epsilon}$$

and

$$\overline{V} = \lim_{\epsilon\to 0} \limsup_{T\to\infty} \overline{V}_{T,\epsilon}$$

are defined similarly. In general, $\overline{V}_{T,\epsilon} \geq \underline{V}_{T,\epsilon}$ and $\overline{V} \geq \underline{V}$ for ADC jamming channels.

Our goal in the information-theoretic framework is to identify $\underline{V}$ and $\overline{V}$ for various model formulations, to understand which encoders and jamming strategies have good performance, and to understand relationships among $\underline{V}$, $\overline{V}$, and C. We find that constructions and bounds obtained in the information-theoretic framework often allow us to say something about actual codes and jamming strategies in the coding framework. However, we have not found a broad class of coding theorems to firmly tie the frameworks together.

We comment briefly on the definitions we have given for the capacity C, and the information rates $\overline{V}$ and $\underline{V}$. Our choices were guided by the theory of arbitrarily varying channels (AVC), initiated by Blackwell, Breiman, and Thomasian [8], as well as by our desire to have $C \leq \overline{V}$. See [12] and [23] for recent surveys of the theory of AVCs, and [34] for a study of the related error exponents. Three aspects of the jammed timing channel do not fit the original formulation of AVCs: i) the jammed timing channel has constraints, namely, an arrival rate constraint on the coder and, in the case of average-delay constraints, a delay constraint on the jammer, ii) the jammed timing channel has memory, and iii) the jammer can causally observe the particular input waveform used. Work of Csiszár and Narayan addresses point i): the use of constraints for AVCs. The paper [13] in particular indicates that in defining the channel capacity C, it is more natural to impose constraints on each codeword and each received word, rather than on averages over codewords or over random codes. This suggests that our definition of channel capacity C is most appropriate. Work of Lapidoth and Telatar [24] addresses point ii): the use of certain AVCs with memory. A

coding theorem is given for which the information rate is defined in a manner similar to our definitions of $\overline{V}$ and $\underline{V}$. Unfortunately, there is no known extension of the AVC theory that addresses point iii).

The original capacity calculations for AVC channels [8] show the benefit of using random codes for AVCs, which is why we allow random codes. Unfortunately, due to point iii) mentioned earlier, Ahlswede's method [2] cannot be used to derandomize the codes, so that implementation would require a source of randomness known to the coder and decoder but not to the jammer. Ahlswede's method can fail even due to point i) when the jammer is constrained if the capacity with an unconstrained jammer is zero [23, p. 2159, top part of column 2].

Blachman [7], Dobrushin [15], McEliece [25], and Hegde *et al.* [18] formulated a jamming game based on mutual information in a different context. McEliece [25] and Hegde *et al.* [18] obtained coding theorems under the assumption that the jammer acts independently on large consecutive time intervals. The paper [14] explores variations of this problem formulation, and shows that the connection between the information game and coding theorems is somewhat tenuous and complex.

However, we find that study of the information game gives substantial insight, if not always proving the existence of codes, for the jammed timing channel.

A. Jamming Channels

In this subsection, we introduce the two generic jamming channels mainly used in this paper.

Two reasonable types of jamming channels are suggested by writing mutual information rate in two different ways (for discrete-time packet waveforms so that entropies are finite)

$$\overline{I}_T(X, W) = \overline{H}_T(X) - \overline{H}_T(X|W_X) = \overline{H}_T(W_X) - \overline{H}_T(W_X|X)$$

with

$$\overline{H}_T(X) = H((X(t), 0 \le t < T))/T$$

the entropy per unit time of process X up to time T. First, a jamming channel can make the output stream random and nearly unrelated to the input so that $\overline{H}_T(X|W_X)$ is large enough that $\overline{H}_T(X|W_X) \approx \overline{H}_T(X)$. Second, a jamming channel can quantize output levels and select regularly spaced times for changes in the output levels so that $\overline{H}_T(W_X)$ is small. We have not found any jamming channels of the first type that perform well, and in fact, a capacity result for the generalized billiard ball channel of T. Berger [5], as discussed in Section V, suggests that jamming channels of this type generally perform poorly. However, jamming channels of the second type perform well in many cases.

Definition II.1: A *periodic dump jammer* with period S for continuous-time packet waveforms is based on a collection of dump times $(\phi, \phi + S, \phi + 2S, \ldots)$ chosen by the jammer independently of the input to the jammer for some $\phi \in (0, S]$. The jammer releases all packets it has at each dump time. (A variation called the fill-alternating periodic dump jammer is defined in Section IV.)

We also make use of a similarly defined periodic dump jammer for discrete-time packet waveforms, where S is an integer and the dump slots are taken to be slots $\phi, \phi + S, \phi + 2S, \ldots$ with ϕ a random variable uniformly distributed on $1, \ldots, S$. Since discrete-time packet waveforms allow only one packet departure per slot, all packets in the jammer's buffer at a dump slot, including a possible packet arriving in the dump slot itself, are transmitted in consecutive slots starting at the dump slot.

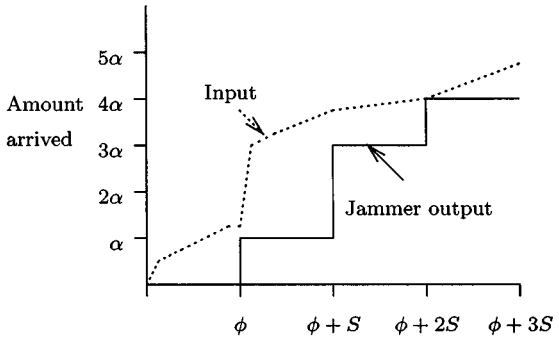

Fig. 4. Periodic quantized dump jammer, $\phi = S$.

The regularly spaced dump times of the periodic dump jammer map many input signals to the same output signals. This makes the output closer to deterministic than the input. Equivalently, it makes it hard to recover the input from the output. Taking this idea one step further—quantizing in both time and number of packets—leads to the following definition.

Definition II.2: A *periodic quantized dump jammer* with parameters S and α for continuous-time fluid waveforms is one that releases $\alpha\lfloor F(t)/\alpha \rfloor$ units of fluid at a dump time t, where $F(t)$ is the amount of fluid in the jammer's buffer at time t and the dump times are taken to be $(\phi, \phi+S, \phi+2S, \ldots)$ for some $\phi \in (0, S]$. (See an example in Fig. 4.)

B. Input Processes

In this subsection, we introduce two generic input processes that perform well for certain classes of jamming channels. We refer to these input processes in the following sections.

Two reasonable types of input processes are suggested by writing mutual information rate in terms of entropy rate (for discrete-time packet waveforms)

$$\overline{I}_T(X, W) = \overline{H}_T(X) - \overline{H}_T(X|W_X).$$

First, an input process can be chosen to be highly random so that $\overline{H}_T(X)$ is large. Second, an input process can be chosen to take advantage of the jammer's delay constraints to ensure that the output is a good predictor of the jammer input so that $\overline{H}_T(X|W_X)$ is small. For a maximum-delay constraint, we have found an input process of the second type that is a max-min optimal input process for continuous-time packet waveforms. There are also input processes of the first type that perform well in certain situations. The following is a class of input processes along the lines of the second type.

Definition II.3: A *batch-with-spacing-S continuous-time packet waveform* is a continuous-time packet waveform such that each jump time is a positive integer multiple of S. That is,

packets are transmitted in batches (the batch sizes remain to be specified) with batches occurring every S seconds.

A *batch-with-spacing-S fluid waveform* is the same, though each batch size can be any positive real number. A *batch-with-spacing-S discrete-time packet waveform* is similar, but a batch can have at most S packets, and the packets of a batch are transmitted in contiguous slots, instead of all at once. A *batch-with-spacing-S rate-constrained fluid packet waveform* is similar, but a batch can have an amount of fluid less than or equal to S, and the fluid of a batch is transmitted at rate one in a time interval.

Another input process of the second type for continuous-time, rate-constrained fluid waveforms and ADC jamming channels is given in Section VI. We obtain a lower bound on capacity for such a jamming channel by using the Gilbert–Varshamov bound to show the existence of a deterministic code with vanishing probability of error.

To obtain an input process of the first type ($\overline{H}_T(X)$ large) for inputs to ADC jamming channels, we would like to consider a batch input for which the batch transmission times are chosen according to a Poisson process. However, such an input does not meet the input constraints for ADC channels. Instead, for a fixed $T > 0$ we consider inputs with a fixed number of packets meeting the ADC input constraints and choose the transmission times of the packets uniformly on $(0, T]$ so the input is similar to a Poisson process.

Definition II.4: For a time $T > 0$, a *Poisson-like input with batch size k* for continuous-time packet waveforms is one which has $k\lfloor \lambda T/k \rfloor$ packets in the interval $(0, T]$, with the packets arranged into batches of size k (k is an integer), and with the transmission times for the batches chosen independently and uniformly on $(0, T]$.

For a fixed time $T > 0$, a *Poisson-like input with batch size S* for continuous-time fluid waveforms is similar, with the modification that S can be real-valued. For N a positive integer multiple of k, a *Bernoulli-like input with batch size k* is a discrete-time packet waveform which has $k\lfloor \lambda N/k \rfloor$ packets in N slots such that packets are grouped into batches of size k with zero or one batches transmitted every super-slot (k slots) and with the batch transmission super-slots chosen uniformly among all $\binom{N/k}{\lfloor \lambda N/k \rfloor}$ possible schemes. Finally, for a fixed time $T > 0$ such that a real-valued S evenly divides T, a version of the *Bernoulli-like input with batch size S* suitable for the rate-constrained fluid model is defined similarly, where batches of fluid are transmitted over randomly selected S-slot windows.

III. Relationships Among Models

In this section, we introduce the relationships among the various packet and fluid models. In addition, we show that $\overline{V} = \underline{V}$ for MDC jammers with discrete-time packet waveforms. Finally, we show that $C \leq \overline{V}$ for each of the jamming channel models.

A. Scaling

We write $\overline{V}_M(\lambda, D)$ to represent the $\min$-$\max$ information per unit time $\overline{V}$ for a rate λ input process and a constrained jammer with delay parameter D and waveforms of type M, where M is one of the following: *CP* for continuous-time packet waveforms, *DP* for discrete-time packet waveforms, *CF* for continuous-time fluid waveforms, and *RF* for continuous-time, rate-constrained fluid waveforms. Similarly, $\underline{V}_M(\lambda, D)$ is the $\max$-$\min$ information per unit time for waveforms of type M. Note that $\overline{V}_M(\lambda, D)$ is taken over all time for model M, whereas $\overline{V}_T$ is defined over $(0, T]$.

Theorem III.1: $\overline{V}_{CP}(\lambda, D)$ and $\underline{V}_{CP}(\lambda, D)$ satisfy the following scaling relationship for MDC or ADC jamming channels:

$$\overline{V}_{CP}(\lambda, D) = \lambda \overline{v}_{CP}(\lambda D)$$

and

$$\underline{V}_{CP}(\lambda, D) = \lambda \underline{v}_{CP}(\lambda D)$$

for some functions $\overline{v}_{CP}(\cdot)$ and $\underline{v}_{CP}(\cdot)$.

Proof of Theorem III.1: Scale time so that one unit of time in the new time scale is $1/\lambda$ units of time on the old time scale. Then on the new scale, the arrival rate is one, the delay constraint is λD, and the information per unit time is $1/\lambda$ times its rate on the old scale. Thus,

$$\overline{V}_{CP}(1, \lambda D) = \overline{V}_{CP}(\lambda, D)/\lambda.$$

So the first relation holds with $\overline{v}_{CP}(\lambda D) = \overline{V}_{CP}(1, \lambda D)$. Similarly, the relationship for $\underline{V}_{CP}$ holds for $\underline{v}_{CP}$ defined analogously. □

Theorem III.2: $\overline{V}_{CF}(\lambda, D)$ and $\underline{V}_{CF}(\lambda, D)$ satisfy the following scaling relationship for MDC or ADC jamming channels:

$$\overline{V}_{CF}(\lambda, D) = \frac{\overline{v}_{CF}}{D}$$

and

$$\underline{V}_{CF}(\lambda, D) = \frac{\underline{v}_{CF}}{D}$$

for some constants $\overline{v}_{CF}$ and $\underline{v}_{CF}$.

Proof of Theorem III.2: Scale time so that one unit of time on the new scale is D units of time on the old scale and scale the size of a unit of fluid so that one unit of fluid on the new scale is equal to λD units of fluid in the old scale. Then on the new scale, the arrival rate is 1, the delay constraint is 1, and the information per unit time is D times its rate on the old scale. Thus,

$$\overline{V}_{CF}(1, 1) = \overline{V}_{CF}(\lambda, D)D.$$

So the first relation holds with $\overline{v}_{CF} = \overline{V}_{CF}(1, 1)$. Similarly, the relation for $\underline{V}_{CF}$ holds for $\underline{v}_{CF}$ defined analogously. □

Theorem III.3: $\overline{V}_{RF}(\lambda, D)$ and $\underline{V}_{RF}(\lambda, D)$ satisfy the following scaling relationship for MDC or ADC jamming channels:

$$\overline{V}_{RF}(\lambda, D) = \frac{\overline{v}_{RF}(\lambda)}{D}$$

and

$$\underline{V}_{RF}(\lambda, D) = \frac{\underline{v}_{RF}(\lambda)}{D}$$

for some functions $\overline{v}_{RF}(\lambda)$ and $\underline{v}_{RF}(\lambda)$.

Proof of Theorem III.3: Scale time so that one unit of time on the new scale is D units of time on the old time scale and scale the size of a unit of fluid so that on the new scale one unit

of fluid is equal to D units of fluid in the old scale. Then on the new scale, the arrival rate is λ units of fluid per unit time, the delay constraint is 1, and the information per unit time is D times its rate on the old scale. Under the new scaling, the arrival rate does not exceed 1 unit of fluid per unit time. Thus,

$$\overline{V}_{RF}(\lambda, 1) = \overline{V}_{RF}(\lambda, D)D.$$

So the first relation holds with $\overline{v}_{RF}(\lambda) = \overline{V}_{RF}(\lambda, 1)$. The relation for $\underline{V}_{RF}$ holds for $\underline{v}_{RF}$ defined similarly. □

B. Relationship Between Packet and Fluid Models

Before introducing the relationships among the fluid models and corresponding packet models, we prove the following useful lemma and state some definitions.

Lemma III.1: Suppose A, $\tilde{A}$, B, $\tilde{B}$ are sets and $f: A\times B \to \Re$ and $\tilde{f}: \tilde{A}\times\tilde{B} \to \Re$ are functions, and suppose for every $\tilde{a} \in \tilde{A}$ there corresponds $a \in A$ and for every $b \in B$ there corresponds $\tilde{b} \in \tilde{B}$ such that $f(a, b) \geq \tilde{f}(\tilde{a}, \tilde{b})$ for all $\tilde{a}$ and b. Then

$$\sup_{a\in A} \inf_{b\in B} f(a, b) \geq \sup_{\tilde{a}\in\tilde{A}} \inf_{\tilde{b}\in\tilde{B}} \tilde{f}\left(\tilde{a}, \tilde{b}\right) \tag{5}$$

and

$$\inf_{b\in B} \sup_{a\in A} f(a, b) \geq \inf_{\tilde{b}\in\tilde{B}} \sup_{\tilde{a}\in\tilde{A}} \tilde{f}\left(\tilde{a}, \tilde{b}\right). \tag{6}$$

Proof of Lemma III.1: Choose any $\tilde{a} \in \tilde{A}$. Then there exists $a \in A$ such that

$$\inf_{b\in B} f(a, b) \geq \inf_{\tilde{b}\in\tilde{B}} \tilde{f}\left(\tilde{a}, \tilde{b}\right) \tag{7}$$

since for each $b \in B$ there exists $\tilde{b} \in \tilde{B}$ such that $f(a, b) \geq \tilde{f}(\tilde{a}, \tilde{b})$. Equation (7) holds for all $\tilde{a} \in \tilde{A}$, so in particular (5) holds. Similarly, choose any $b \in B$. Then there exists $\tilde{b} \in \tilde{B}$ such that

$$\sup_{a\in A} f(a, b) \geq \sup_{\tilde{a}\in\tilde{A}} \tilde{f}\left(\tilde{a}, \tilde{b}\right) \tag{8}$$

since for each $\tilde{a} \in \tilde{A}$ there exists $a \in A$ such that $f(a, b) \geq \tilde{f}(\tilde{a}, \tilde{b})$. Equation (8) holds for all $b \in B$, so in particular (6) holds. □

Definition III.1: A *trigger packetizer* is a device that converts a fluid waveform $A(t)$ into the continuous time packet waveform $\lceil A(t) \rceil$. (A trigger packetizer in isolation is not physically realizable because it advances partial packets. See illustration in Fig. 5.)

Definition III.2: An *accumulate-and-dump packetizer* is a device that converts a fluid waveform $A(t)$ into the continuous-time packet waveform $\lfloor A(t) \rfloor$. (See illustration in Fig. 6.)

Definition III.3: A *fluidizer* is a device which converts any discrete-time packet waveform into a continuous —time, rate-constrained waveform. A packet arriving at slot k is replaced by fluid flow with rate one on the interval $[k, k+1)$, as shown in Fig. 7.

Definition III.4: A *slotter* is a device which converts a continuous-time packet waveform having no arrivals in $[0, 1)$ into a discrete-time packet waveform. An arrival at the device at time t departs in slot $\lfloor t \rfloor$, or the first available slot after that if previous packets are waiting to be transmitted. The slotter is illustrated in Fig. 8. (A slotter in isolation is not a physically realizable device since packets can depart before they arrive).

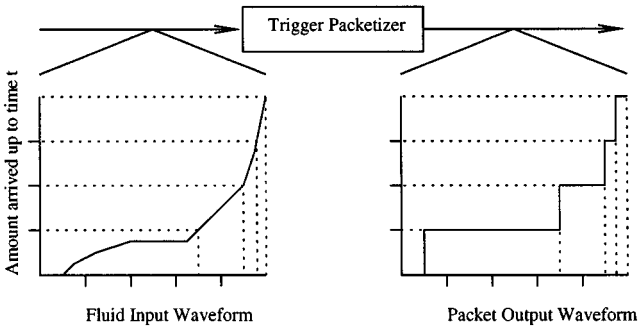

Fig. 5. Trigger packetizer.

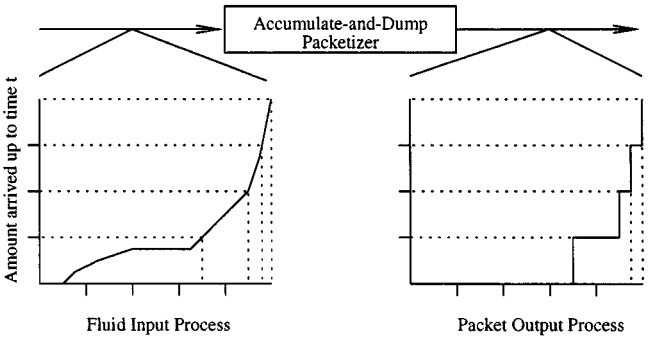

Fig. 6. Accumulate-and-dump packetizer.

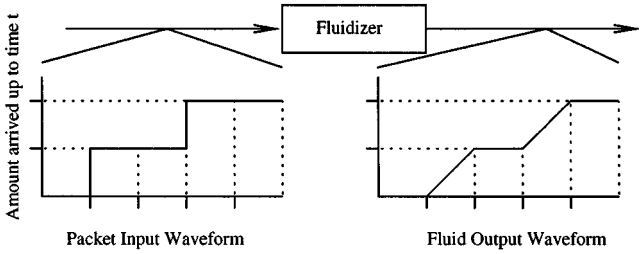

Fig. 7. Fluidizer.

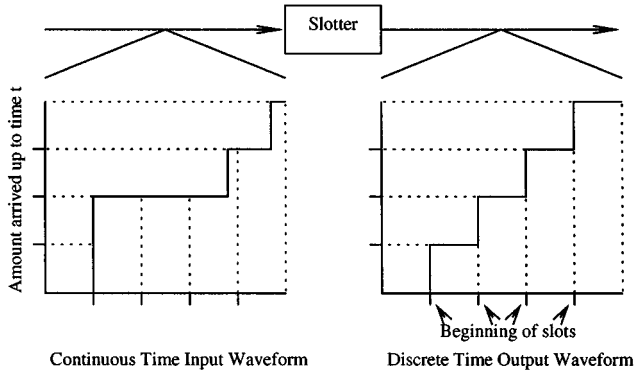

Fig. 8. Slotter.

We use these definitions and Lemma III.1 to show relationships among the various channel models. First we describe the relationships between the continuous-time packet models and

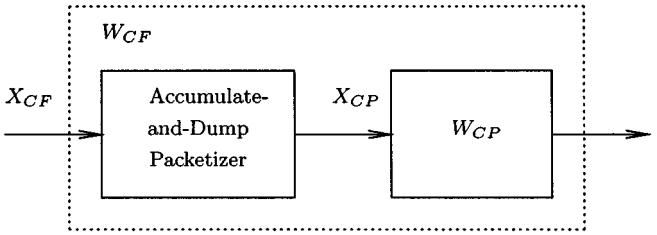

Fig. 9. Relationships of transformations, $CP \geq CF$ with extra delay.

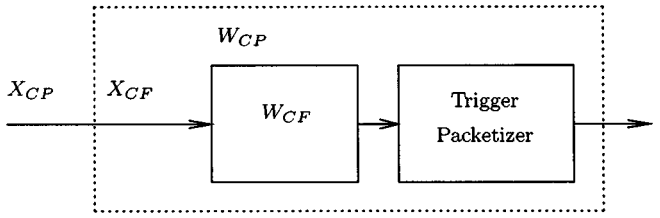

Fig. 10. Relationships of transformations, $CF \geq CP$.

the continuous-time fluid models. Next, we describe the relationships between the discrete-time packet models and the continuous-time, rate-constrained fluid models.

Theorem III.4: For ADC jamming channels

$$\overline{V}_{CP}(\lambda, D) \geq \overline{V}_{CF}\left(\lambda, D\left(1+\frac{1}{\lambda D}\right)\right)$$

and

$$\underline{V}_{CP}(\lambda, D) \geq \underline{V}_{CF}\left(\lambda, D\left(1+\frac{1}{\lambda D}\right)\right).$$

Proof of Theorem III.4: For an interval of length $T > 0$ and ϵ such that $0 < \epsilon < \lambda$, let $\mathcal{X}_{\mathcal{CP}}$ (short for $\mathcal{X}_{T,\epsilon}$ for continuous-time packet waveforms) be the set of all rate $\lambda \pm \epsilon$ input processes for ADC jamming channels for continuous-time packet waveforms, let $\mathcal{X}_{CF}$ be the set of all rate $\lambda \pm \epsilon$ input processes for ADC jamming channels for continuous-time fluid waveforms, let $\mathcal{W}_{CP}(D)$ be the set of all ADC jamming channels with delay parameter D for continuous-time packet waveforms, and let $\mathcal{W}_{CF}(D)$ be the set of all ADC jamming channels with delay parameter D for continuous-time fluid waveforms.

For every packet jamming channel $W_{CP} \in \mathcal{W}_{CP}(D)$, there exists a fluid jamming channel $W_{CF} \in \mathcal{W}_{CF}(D(1+\frac{1}{(\lambda-\epsilon)D}))$ obtained by concatenating an accumulate-and-dump packetizer with the packet jamming channel W_{CP} as shown in Fig. 9. The resulting fluid jammer is ADC with delay parameter at most $D(1+\frac{1}{(\lambda-\epsilon)D})$ since the accumulate-and-dump packetizer holds at most one unit of fluid at a time and hence, by Little's law, introduces a mean delay of at most $1/(\lambda-\epsilon)$. Similarly, for every rate λ fluid input process $X_{CF} \in \mathcal{X}_{CF}$, there exists a rate λ packet input process $X_{CP} \in \mathcal{X}_{CP}$ obtained by passing X_{CF} through an accumulate-and-dump packetizer.

Using the data processing inequality of information theory (cf. [10, p. 32]), it is easy to see that for any fixed $T > 0$

$$\overline{I}_T(X_{CP}, W_{CP}) \geq \overline{I}_T(X_{CF}, W_{CF}).$$

Then applying Lemma III.1 with A representing $\mathcal{X}_{CP}$, $\tilde{A}$ representing $\mathcal{X}_{CF}$, B representing $\mathcal{W}_{CP}(D)$, $\tilde{B}$ representing $\mathcal{W}_{CF}(D(1+\frac{1}{(\lambda-\epsilon)D}))$, and f and $\tilde{f}$ representing $\overline{I}_T$, and taking appropriate limits, the result is shown. □

Theorem III.5: For MDC, MBC, and ADC jamming channels

$$\overline{V}_{CF}(\lambda, D) \geq \overline{V}_{CP}(\lambda, D)$$

$$\underline{V}_{CF}(\lambda, D) \geq \underline{V}_{CP}(\lambda, D).$$

Proof of Theorem III.5: To be specific, we show the result for MDC jamming channels, but essentially the same proof works for MBC and ADC jamming channels. For an interval of length $T > 0$, let $\mathcal{X}_{CP}$ be the set of all rate λ input processes for continuous-time packet waveforms, let $\mathcal{X}_{CF}$ be the set of all rate λ input processes for continuous-time fluid waveforms, let $\mathcal{W}_{CP}(D)$ be the set of all MDC jamming channels with delay parameter D for continuous-time packet waveforms, and let $\mathcal{W}_{CF}(D)$ be the set of all MDC jamming channels with delay parameter D for continuous-time fluid waveforms.

Given a fluid jamming channel $W_{CF} \in \mathcal{W}_{CF}(D)$, there exists a packet jamming channel $W_{CP} \in \mathcal{W}_{CP}(D)$, obtained by concatenating the fluid jamming channel W_{CF} with a trigger packetizer as in Fig. 10. Since $\mathcal{X}_{CP} \subset \mathcal{X}_{CF}$, the fluid jammer will meet its delay constraints for any packet input process. In addition, the trigger packetizer advances partial packets and does not add any delay. While trigger packetizers in general are not physically realizable, the resulting concatenation of W_{CF} with a trigger packetizer is realizable for continuous-time packet inputs. Thus, the concatenation of a fluid jammer with delay parameter D and a trigger packetizer results in a packet jammer with delay parameter no larger than D. Also, for every rate λ packet input process $X_{CP} \in \mathcal{X}_{CP}$, there is a rate λ fluid input process $X_{CF} \in \mathcal{X}_{CF}$, namely, $X_{CF} = X_{CP}$.

Applying the data processing inequality (cf. [10, p. 32]) it is easy to see that for any fixed time $T > 0$, $\overline{I}_T(X_{CF}, W_{CF}) \geq \overline{I}_T(X_{CP}, W_{CP})$. Then applying Lemma III.1 with A representing $\mathcal{X}_{CF}$, $\tilde{A}$ representing $\mathcal{X}_{CP}$, B representing $\mathcal{W}_{CF}(D)$, $\tilde{B}$ representing $\mathcal{W}_{CP}(D)$, and f and $\tilde{f}$ representing $\overline{I}_T$, and taking appropriate limits, the result is shown. □

Corollary III.1: For ADC jamming channels as $\lambda D \to \infty$

$$\overline{V}_{CP}/\overline{V}_{CF} \to 1 \quad \text{and} \quad \underline{V}_{CP}/\underline{V}_{CF} \to 1.$$

Proof of Corollary III.1: By the scaling results of Theorem III.2 we have that for all λ and D, $\overline{V}_{CF}(\lambda, D) = \frac{\overline{v}_{CF}}{D}$ for some constant $\overline{v}_{CF}$ and $\underline{V}_{CF}(\lambda, D) = \frac{\underline{v}_{CF}}{D}$ for some constant $\underline{v}_{CF}$. Applying the results of Theorems III.4 and III.5 for ADC jamming channels, we have that

$$\overline{V}_{CF} = \frac{\overline{v}_{CF}}{D} \geq \overline{V}_{CP}(\lambda, D) \geq \frac{\overline{v}_{CF}}{D\left(1+\frac{1}{\lambda D}\right)} = \frac{\overline{V}_{CF}}{\left(1+\frac{1}{\lambda D}\right)}$$

and

$$\underline{V}_{CF} = \frac{\underline{v}_{CF}}{D} \geq \underline{V}_{CP}(\lambda, D) \geq \frac{\underline{v}_{CF}}{D\left(1+\frac{1}{\lambda D}\right)} = \frac{\underline{V}_{CF}}{\left(1+\frac{1}{\lambda D}\right)}$$

for all λ and D. Thus, $\overline{V}_{CP}/\overline{V}_{CF}$ and $\underline{V}_{CP}/\underline{V}_{CF}$ are in the interval $[(1+\frac{1}{\lambda D})^{-1}, 1]$ for all λ and D, and the result follows. □

Theorem III.6: For ADC jamming channels we have

$$\overline{V}_{DP}(\lambda, D) \geq \overline{V}_{RF}\left(\lambda, D\left(1+\frac{1}{\lambda D}+\frac{1}{2D}\right)\right)$$

and

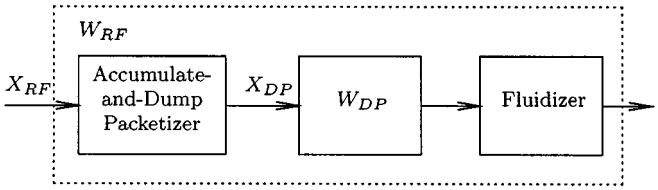

Fig. 11. Relationships of transformations, $DP \geq RF$ with extra delay.

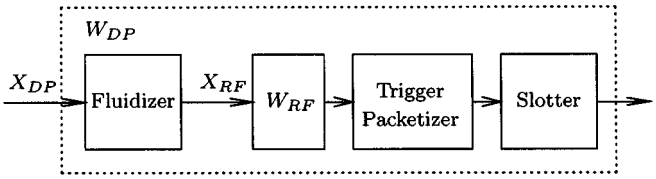

Fig. 12. Relationships of transformations, $RF \geq DP$.

$$\underline{V}_{DP}(\lambda, D) \geq \underline{V}_{RF}\left(\lambda, D\left(1+\frac{1}{\lambda D}+\frac{1}{2D}\right)\right).$$

Proof of Theorem III.6: For an interval of length $T > 0$ and ϵ such that $0 < \epsilon < \lambda$, let $\mathcal{X}_{DP}$ be the set of all rate λ input processes for discrete-time packet waveforms, let $\mathcal{X}_{RF}$ be the set of all rate λ input processes for continuous-time, rate-constrained fluid waveforms, let $\mathcal{W}_{DP}(D)$ be the set of all ADC jamming channels with delay parameter D for discrete-time packet waveforms, and let $\mathcal{W}_{RF}(D)$ be the set of all ADC jamming channels with delay parameter D for continuous-time, rate-constrained fluid waveforms.

For every rate λ continuous-time, rate-constrained fluid input process $X_{RF} \in \mathcal{X}_{RF}$, there exists a rate-$\lambda$ discrete packet input process $X_{DP} \in \mathcal{X}_{DP}$ obtained by passing X_{RF} through an accumulate-and-dump packetizer. For every packet jamming channel $W_{DP} \in \mathcal{W}_{DP}(D)$, there exists a fluid jamming channel

$$W_{RF} \in \mathcal{W}_{RF}\left(D\left(1+\frac{1}{(\lambda-\epsilon)D}+\frac{1}{2D}\right)\right)$$

obtained by concatenating an accumulate-and-dump packetizer with the packet jamming channel W_{CP} and a fluidizer as in Fig. 11. The resulting fluid jammer is a continuous-time, rate-constrained ADC jammer with delay parameter $D(1+\frac{1}{(\lambda-\epsilon)D}+\frac{1}{2D})$. To see this note that any continuous-time, rate-constrained fluid waveform is converted to a discrete-time packet waveform by the accumulate-and-dump packetizer and the accumulate-and-dump packetizer holds at most one unit of fluid at a time and hence introduces a mean delay of at most $1/(\lambda-\epsilon)$. The fluidizer introduces a mean delay of at most $1/2$ slot times since the packet is released at rate 1 during one slot time.

Applying the data processing inequality [10, p. 32], it is easy to see that for any fixed time $T > 0$, $\overline{I}_T(X_{DP}, W_{DP}) \geq \overline{I}_T(X_{RF}, W_{RF})$. Then applying Lemma III.1 with A representing $\mathcal{X}_{DP}$, $\tilde{A}$ representing $\mathcal{X}_{RF}$, B representing $\mathcal{W}_{DP}(D)$, $\tilde{B}$ representing $\mathcal{W}_{RF}(D(1+\frac{1}{(\lambda-\epsilon)D}+\frac{1}{2D}))$, and f and $\tilde{f}$ representing $\overline{I}_T$, and taking appropriate limits, the result is shown. □

Theorem III.7: For MDC, MBC, and ADC jamming channels we have

$$\overline{V}_{RF}(\lambda, D) \geq \overline{V}_{DP}(\lambda, D)$$

and

$$\underline{V}_{RF}(\lambda, D) \geq \underline{V}_{DP}(\lambda, D).$$

Proof of Theorem III.7: To be specific, we show the result for MDC jamming channels, but the proof holds equally for MBC and ADC jamming channels. For an interval of length $T > 0$, let $\mathcal{X}_{DP}$ be the set of all rate-λ input processes for discrete-time packet waveforms, let $\mathcal{X}_{RF}$ be the set of all rate-λ input processes for continuous-time, rate-constrained fluid waveforms, let $\mathcal{W}_{DP}(D)$ be the set of all MDC jamming channels with delay parameter D for discrete-time packet waveforms, and let $\mathcal{W}_{RF}(D)$ be the set of all MDC jamming channels with delay parameter D for continuous-time, rate-constrained fluid waveforms.

For every rate λ packet input process $X_{DP} \in \mathcal{X}_{DP}$, there is a rate λ fluid input process $X_{RF} \in \mathcal{X}_{RF}$ obtained by passing X_{DP} through a fluidizer. For every fluid jamming channel $W_{RF} \in \mathcal{W}_{RF}(D)$, there exists a packet jamming channel $W_{DP} \in \mathcal{W}_{DP}(D)$, obtained by concatenating a fluidizer, the fluid jamming channel W_{RF}, a trigger packetizer, and a slotter as shown in Fig. 12. The fluidizer adds exactly $1/2$ unit of delay to each packet, while the trigger packetizer subtracts at least $1/2$ unit of delay from each packet-sized unit of fluid. Taken together, these devices add no additional delay to a discrete-time packet waveform. For a discrete-time packet input, the output of the trigger packetizer will have at most one packet per slot since the jammer W_{RF} enforces at most one unit of fluid per unit time and the input to the slotter will have no packets in $[0, 1)$ since $X_{DP} \in \mathcal{X}_{DP}$. Thus, the slotter only advances packets to the nearest slot time, so it does not add any delay. The only delay in the resulting discrete-time packet jammer is that introduced by fluid jammer and the resulting delay parameter will be no larger in the discrete-time packet jammer than in the continuous-time, rate-constrained fluid jammer.

As before, for any fixed $T > 0$

$$\overline{I}_T(X_{RF}, W_{RF}) \geq \overline{I}_T(X_{DP}, W_{DP}).$$

Then applying Lemma III.1 with A representing $\mathcal{X}_{RF}$, $\tilde{A}$ representing $\mathcal{X}_{DP}$, B representing $\mathcal{W}_{RF}(D)$, $\tilde{B}$ representing $\mathcal{W}_{DP}(D)$, and f and $\tilde{f}$ representing $\overline{I}_T$, and taking appropriate limits, the result is shown. □

Corollary III.2: For ADC jamming channels as $D \to \infty$ and $\lambda D \to \infty$ (for example, with λ fixed)

$$\overline{V}_{DP}/\overline{V}_{RF} \to 1 \quad \text{and} \quad \underline{V}_{DP}/\underline{V}_{RF} \to 1.$$

Proof of Corollary III.2: The proof is similar to that of Corollary III.1. Theorems III.3, III.6, and III.7 imply that $\overline{V}_{DP}/\overline{V}_{RF}$ and $\underline{V}_{DP}/\underline{V}_{RF}$ are in the interval $[(1+\frac{1}{\lambda D}+\frac{1}{2D})^{-1}, 1]$ for all λ and D, which concludes the proof. □

C. Saddle-Point Existence for MDC Jammer for Discrete-Time Packet Waveforms

In this subsection, we show that $\overline{V} = \underline{V}$ for MDC jamming channels for discrete-time packet waveforms (i.e., $\overline{V}_{DP} = \underline{V}_{DP}$). First we prove two useful lemmas.

Lemma III.2: For MDC jamming channels with delay parameter D and discrete-time packet waveforms, $\overline{V}_T = \underline{V}_T$ for all positive integers T.

Proof of Lemma III.2: Fix T a positive integer. A discrete-time packet waveform on the interval $[0, T]$ can be represented by a vector in $\{0, 1\}^T$, such that the ith coordinate is one if a packet arrives in slot i. Thus, $\mathcal{X}_T$ can be viewed as the set of probability distributions on $\{0, 1\}^T$ such that the mean number of arrivals satisfies $E[x_1 + \cdots + x_T] = \lambda T$. A probability distribution on $\{0, 1\}^T$ can be expressed as $(p(x)\colon x \in \{0, 1\}^T)$, which is an element of $\Re^{\{0,1\}^T}$, the space of all vectors with index set $\{0, 1\}^T$. Moreover, $\mathcal{X}_T$ is a closed, bounded, convex subset of $\Re^{\{0,1\}^T}$. Similarly, let $\mathcal{W}$ represent the set of conditional probability density functions $\{q(y|x)\colon x, y \in \{0, 1\}^T\}$ satisfying the causality constraints

$$q(y_1, \ldots, y_k|x) = q(y_1, \ldots, y_k|\tilde{x})$$

if $x_1 = \tilde{x}_1, \ldots, x_k = \tilde{x}_k$, and satisfying

$$q(y_1, \ldots, y_n|x_1, \ldots, x_n) = 0$$

if x and y violate the maximum delay less than D constraint. Then $\mathcal{W}$ is a closed, bounded, convex subset of $\Re_+^{\{0,1\}^{2T}}$.

Since mutual information $I(X;Y)$ is a concave function of $p(x)$ for fixed $p(y|x)$ and a convex function of $p(y|x)$ for (X, Y) distributed according to $p(x, y) = p(x)p(y|x)$, the classical min-max theorem (e.g., [22]) implies that

$$\min_{\mathcal{W}} \max_{\mathcal{X}_T} \overline{I}_T(X, W) = \max_{\mathcal{X}_T} \min_{\mathcal{W}} \overline{I}_T(X, W)$$

and the result follows. □

Lemma III.3: Given a nondecreasing function f on $\Re_+$ or $\mathcal{Z}_+$ and some $b > 0$ such that $f(s + t + b) \geq f(s) + f(t)$ for all $s, t > 0$, the limit $\lim_{t\to\infty} f(t)/t$ exists.

Proof of Lemma III.3: Let $F(t) = f(t - b)$ and let $\tilde{s} = s + b$ and $\tilde{t} = t + b$. Then

$$F(\tilde{s} + \tilde{t}) \geq F(\tilde{s}) + F(\tilde{t}), \qquad \text{for } \tilde{s}, \tilde{t} > b$$

which in turn implies that $\frac{F(n\tilde{t})}{n\tilde{t}} \geq \frac{F(\tilde{t})}{\tilde{t}}$ for all n and for all $\tilde{t} > b$.

Let $L = \limsup_{t\to\infty} F(t)/t$ and let $L' < L$. Select $B > b$ such that $F(B)/B \geq L'$. Then

$$\frac{F(nB)}{nB} \geq \frac{F(B)}{B} \geq L'$$

for all n. Since F is nondecreasing, this implies that

$$\liminf_{t\to\infty} \frac{F(t)}{t} \geq L'.$$

Therefore, $\lim_{t\to\infty} \frac{F(t)}{t}$ exists, which in turn implies that $\lim_{t\to\infty} \frac{f(t)}{t}$ exists. □

Theorem III.8: For MDC jamming channels with delay parameter D and discrete-time packet waveforms, $\overline{V} = \underline{V}$ and a mutual information rate saddle point exists.

Proof of Theorem III.8: For $T > 0$, let $\mathcal{X}_T$ be the set of rate λ input processes for discrete-time packet waveforms and let $\mathcal{W}$ be the set of MDC jammers for discrete-time packet waveforms. By Lemma III.2, we have that $\overline{V}_T = \underline{V}_T$ for all T.

Take

$$f(t) = t\underline{V}_t = \max_{X\in\mathcal{X}_t} \min_{W\in\mathcal{W}} I_t(X; W_X).$$

Then f is clearly nondecreasing. Fix integers $s, t > 0$. Let

$$X^t = \arg\max_{X\in\mathcal{X}_t} \min_{W\in\mathcal{W}} I_t(X; W_X)$$

be a maximizing input for $f(t)$ and, similarly, let X^s be a maximizing input for $f(s)$.

Construct a new input X^* which consists of X^s for the first s time units, a transition for the next $b = \lceil \frac{D}{1-\lambda} \rceil$ time units (described next), and a translation of X^t for the last t time units. Each of the first $\lfloor \lambda b \rfloor$ slots of the transition interval have an arriving packet, the next slot has a packet with probability $\lambda b - \lfloor \lambda b \rfloor$, and all of the remaining slots of the transition interval have no packets. The mean arrival rate during the transition interval, and hence for the whole input, is λ. Since

$$b - (\lfloor \lambda b \rfloor + 1) \geq (1 - \lambda)b - 1 \geq D - 1$$

there are no packets input for at least $D - 1$ slots at the end of the transition interval. Thus, there are no packets left in the jammer just after the first $s+b$ slots. Using the notation $X^*_{(a,b]} = (X^*(t)\colon a+1 \leq t \leq b)$, and letting $q = s+b$ and $r = s+b+t$

$$f(r) \geq \min_{W\in\mathcal{W}} I(X^*; W_{X^*}) \tag{9}$$

$$= \min_{W\in\mathcal{W}} \begin{pmatrix} I\left(X^*_{(0,q]}; W_{X^*}\right) \\ +I\left(X^*_{(q,r]}; W_{X^*} \middle| X^*_{(0,q]}\right) \end{pmatrix} \tag{10}$$

$$\geq \min_{W\in\mathcal{W}} \begin{pmatrix} I\left(X^s; W_{X^s}\right) \\ +H\left(X^*_{(q,r]} \middle| X^*_{(0,q]}\right) \\ -H\left(X^*_{(q,r]} \middle| X^*_{(0,q]}, W_{X^*}\right) \end{pmatrix} \tag{11}$$

$$\geq \min_{W\in\mathcal{W}} \begin{pmatrix} I(X^s; W_{X^s}) \\ +H\left(X^*_{(q,r]}\right) \\ -H\left(X^*_{(q,r]} \middle| W_{X^*}\right) \end{pmatrix} \tag{12}$$

$$\geq \min_{W\in\mathcal{W}} I(X^s; W_{X^s}) + \min_{W\in\mathcal{W}} I(X^t; W_{X^t}) \tag{13}$$

$$\geq f(s) + f(t) \tag{14}$$

where (12) follows since conditioning only decreases entropy and since $X^*_{(s+b,\,s+b+t]}$ is independent of $X^*_{(0,\,s+b]}$; and (13) follows since the transition interval empties the jammer's

buffer. Thus, by Lemma III.3, $\lim_{t\to\infty} f(t)/t$ exists and hence $\lim_{t\to\infty} \underline{V}_t$ exists. Therefore, $\overline{V}_{DP} = \underline{V}_{DP}$. □

D. Relationship Between Capacity and $\overline{V}$

In this subsection, we show that in general $C \le \overline{V}$.

Theorem III.9: The capacity C for a class of MDC, MBC, or ADC jammers satisfies $C \le \overline{V}$, where codewords are taken to be rate λ codewords, and input processes are rate λ input processes.

Proof of Theorem III.9: The proof is similar to the standard converse coding theorem for channel capacity. Let $T > 0$ and $\epsilon > 0$. Let $\mathcal{W}$ be the class of MDC, MBC, or ADC jamming channels. For every $R < C$ and $\epsilon > 0$, there exists a sequence of random (M, T) codes with $\frac{1}{T}\log M \ge R - \epsilon$, with

$$P\left(\left|\frac{X_m(T)}{T} - \lambda\right| \le \epsilon\right) = 1, \qquad \text{for } 1 \le m \le M$$

and with maximum probability of error over all messages m and over all jammers in $\mathcal{W}$ converging to zero as $T \to \infty$.

Since the maximal probability of error converges to zero, the average probability of error $P_e^{(T)}$ averaged over all messages for any $W \in \mathcal{W}$ also converges to zero as $T \to \infty$. For a given T, let m be a message index uniformly distributed over $\{1, \ldots, M\}$, and let θ be a random variable, independent of m, representing the choice of codebook. Consider the input consisting of a codeword depending on the random variables m and θ, $X_{m,\theta}$. Fix $W \in \mathcal{W}$ and let Y represent the corresponding output waveform for jammer W up to time T. Using the independence of m and θ

$$\log M = H(m|\theta) = H(m|Y, \theta) + I(m; Y|\theta) \tag{15}$$

$$\le H(m|Y, \theta) + \sum_{\theta_0 \in \theta} I(X_{m,\theta_0}; Y|\theta = \theta_0)\mathrm{p}(\theta = \theta_0) \tag{16}$$

$$\le \log 2 + P_e^{(T)} \log M + \max_{\theta_0 \in \mathbf{\Theta}} I(X_{m,\theta_0}; Y|\theta = \theta_0) \tag{17}$$

where (16) follows from the data processing inequality and (17) follows from Fano's inequality. Since (17) holds for each $W \in \mathcal{W}$, we have that

$$R \le \epsilon + \frac{1}{T}\log M$$
$$\le \frac{\log 2}{(1 - P_e^{(T)})T} + \inf_{W\in\mathcal{W}} \max_{\theta_0\in\mathbf{\Theta}} \frac{I(X_{m,\theta_0}; Y|\theta = \theta_0)}{(1 - P_e^{(T)})T}.$$

Note that $X_{m,\theta} \in \mathcal{X}_{T,\epsilon}$ for the ADC model. For the MDC or MBC model, choose a slightly larger time interval $\tilde{T} > T$ and add to each codeword an appropriate number of packets (or fluid) in the interval $(T, \tilde{T}]$ to obtain codewords $\tilde{X}_{m,\theta}$ so the random input satisfies $E[\tilde{X}_m(\tilde{T})] = \lambda\tilde{T}$. Taking limits we get that $R \le \overline{V}$ for each $R < C$ so that the result follows. □

IV. MDC Jammers

In this section, we present bounds for $\overline{V}$ and $\underline{V}$ with rate λ input processes and MDC jammers. We also show that the capacity of an MDC jammer for continuous-time packet waveforms is equal to the saddle-point information rate and that codes exist to achieve the rate even when the decoder does not have access to timing information. The four waveform types are considered in the same order in this section and in each of the next two sections.

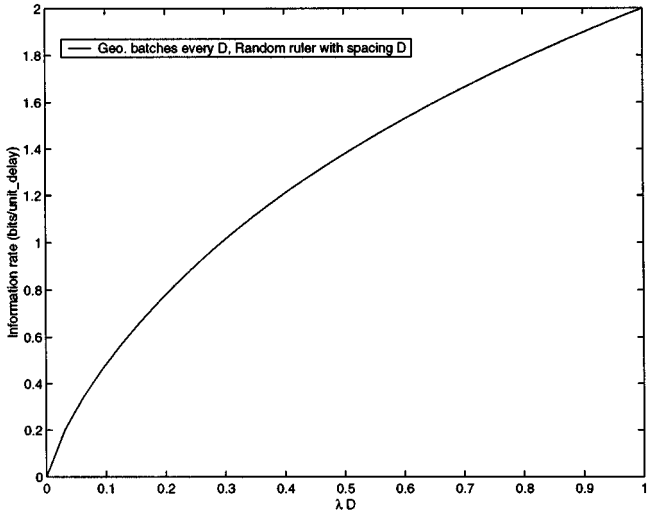

Fig. 13. Value of MDC game for continuous-time packet waveforms.

A. MDC, Continuous-Time Fluid Model

For continuous-time fluid waveforms, the capacity of an MDC jammer with delay parameter D is infinite. This result can be seen by considering a rate λ code with batches every D time units where the batch size is a positive, real-valued random variable with mean λD. Since MDC jammers cannot delay fluid by D or more time units, the batch size can be communicated error-free through any MDC jammer. Since a positive real-valued random variable can convey an infinite number of bits, $C = \overline{V}_{CF} = \underline{V}_{CF} = \infty$.

B. MDC, Continuous-Time Packet Model

Given $\mu > 0$, let $\mathrm{Geo}_0(\mu)$ represent the probability distribution on the nonnegative integers given by $p_k = \eta(1-\eta)^k$, where $\eta = \frac{1}{1+\mu}$. This distribution has maximum entropy among all distributions on the nonnegative integers with mean μ. Let X^* be a batch-with-spacing-D input process such that each batch size is chosen independently with distribution $\mathrm{Geo}_0(\lambda D)$. We call X^* a $\mathrm{Geo}_0(\lambda D)$-batch-with-spacing-$D$ input process. Let W^* denote a periodic dump jammer with period D.

Theorem IV.1: For MDC jamming channels with delay parameter D and continuous-time packet waveforms, (X^*, W^*) is a saddle point, and the saddle-point information rate is $\frac{H(\mathrm{Geo}_0(\lambda D))}{D}$ (see Fig. 13). Hence X^* is max-min optimal and the periodic dump jammer is min-max optimal.

Proof of Theorem IV.1: It suffices to show that if $\mathcal{W}$ is the set of MDC jammers with delay parameter D and $\mathcal{X}_T$, $T > 0$ denotes the rate λ input processes for $[0, T]$, then

$$\limsup_{T\to\infty} \sup_{X\in\mathcal{X}_T} \overline{I}_T(X, W_X^*) \le \frac{1}{D} H(\mathrm{Geo}_0(\lambda D)) \tag{18}$$

and

$$\liminf_{T\to\infty} \inf_{W\in\mathcal{W}} \overline{I}_T(X^*, W_{X^*}) \geq \frac{1}{D} H(\mathrm{Geo}_0(\lambda D)). \quad (19)$$

First we show (18). For any $T > 0$ which is an integer multiple of D, and a fixed $X \in \mathcal{X}_T$

$$I_T(X; W_X^*) \leq I_T(X; W_X^*, \phi) = I_T(X; W_X^* | \phi) \quad (20)$$

$$= H((W_X^*(t), 0 \leq t \leq T) | \phi) \quad (21)$$

$$\leq \sum_{i=1}^{\frac{T}{D}} H(Y_i) \quad (22)$$

where ϕ is the initial dump time and Y_i is the number of packets that depart at the ith dump time. Equation (21) follows because under jammer W^*, the output is completely determined by X and ϕ, and (22) follows since conditioning only reduces entropy.

Let $\mu_i = E[Y_i]$ for each i. Then with no other constraints we have that $H(Y_i) \leq H(\mathrm{Geo}_0(\mu_i))$ by the maximum entropy property of the $\mathrm{Geo}_0(\mu)$ distribution. Let $n = \frac{T}{D}$. The average number of packets that arrive at the jammer up to time T is λT, so $\sum_{i=1}^n \mu_i$, the mean number of packets that depart the jammer up to time T, is at most λT. Thus, $\frac{1}{n}\sum_{i=1}^n \mu_i \leq \lambda D$. It can easily be seen that $H(\mathrm{Geo}_0(x))$ is convex and monotone increasing in x, so by Jensen's inequality

$$I_T(X; W_X^*) \leq \sum_{i=1}^n H(\mathrm{Geo}_0(\mu_i)) \leq nH(\mathrm{Geo}_0(\lambda D)).$$

Thus, $\overline{I}_T(X, W_X^*) \leq H(\mathrm{Geo}_0(\lambda D))/D$ for all $X \in \mathcal{X}_T$ and (18) holds.

On the other hand, note that for any $W \in \mathcal{W}$ and $T > D$

$$I_T(X^*; W_{X^*}) \geq H((X^*(t), 0 \leq t \leq T - D)) - H\begin{pmatrix} (X^*(t), 0 \leq t \leq T-D) | \\ (W_{X^*}(t), 0 \leq t \leq T) \end{pmatrix}$$

$$= H((X^*(t), 0 \leq t \leq T - D)) \quad (23)$$

$$= \left\lfloor \frac{T-D}{D} \right\rfloor H(\mathrm{Geo}_0(\lambda D)) \quad (24)$$

where (23) follows from the fact that for any jammer delaying packets by strictly less than D time slots, the input X^* up to time $T - D$ is completely determined by the output up to time T. Equation (24) follows because X^* has $\lfloor \frac{T-D}{D} \rfloor$ independent batches of arrivals by time $T - D$. Therefore, (19) is true, and Theorem IV.1 has been shown. □

Next, we show that the capacity of an MDC jammer for continuous-time packet waveforms is equal to the saddle-point information rate. Moreover, we show the existence of codes that achieve the saddle-point information rate, even when the decoder does not have access to timing information.

Consider a block timing code with codewords of duration DN time units. Each codeword is such that packets are transmitted only at times which are a positive integer multiple of D, so that a codeword can be identified with the vector of batch sizes $(x_1, \ldots, x_N)$. The code will be taken to be the set of all $(x_1, \ldots, x_N)$ with nonnegative integer coordinates such that $x_1 \geq 1$, $x_N = 0$, $\sum_{j=1}^N x_j = \lceil \lambda DN \rceil$, and $\sum_{j=1}^N j x_j = s$,

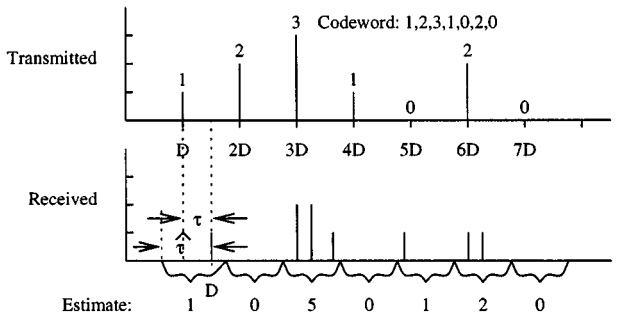

Fig. 14. Coding and decoding: decoder does not know time that coder begins transmitting.

where s is a fixed constant specified below. The sum $\sum_{j=1}^N j x_j$ is called the *delay sum* for the sequence $(x_1, \ldots, x_N)$.

We first show that for such codewords, error-free decisions can be made at the decoder, even without timing information. Next, we specify the delay sum s and show that there are enough codewords in this code to achieve the saddle-point information rate.

The decoder works as follows. Let time D for the coder be the time that the first batch is transmitted and let time D for the decoder be the time that the first packet arrives. To determine the transmitted batch sizes $X_1, \ldots, X_N$, the decoder would like to know τ, the delay of the first packet received, as in Fig. 14. By guessing a delay $\hat{\tau}$, the decoder can get for each j an estimate of X_j, denoted $\hat{X}_j$, by counting packets in the interval $[jD - \hat{\tau}, (j+1)D - \hat{\tau})$, as shown in Fig. 14. The codeword in Fig. 14 with $N = 7$ satisfies $\sum_{i=1}^N i X_i = s = 30$. Note that if $\hat{\tau} = \tau$, the estimate of the codeword equals the codeword, and, of course, the estimate of the codeword has the same delay sum s as the codeword. As a function of $\hat{\tau}$, the delay sum of the estimated codeword is nondecreasing and its jumps mark changes in the estimated codeword. Thus, the decoder can correctly find the codeword by adjusting $\hat{\tau}$ until $\sum_{j=1}^N j\hat{X}_j = s$. For example, for the $\hat{\tau}$ shown in Fig. 14, the delay sum of the codeword estimate is 33, indicating that $\hat{\tau}$ is too large. Therefore, a codeword can be transmitted without error every DN time units.

Next we specify the delay sum s and show that there are enough codewords to achieve the saddle-point information rate. Let $\mathcal{U}$ be the collection of sequences of nonnegative integers $(x_1 \ldots, x_N)$, such that $x_1 \geq 1$, $x_N = 0$, and $\sum_{j=1}^N x_j = \lceil \lambda DN \rceil$. Then

$$|\mathcal{U}| = \binom{\lceil \lambda DN \rceil + N - 3}{N-2} \quad (25)$$

$$\geq \frac{2^{(\lceil \lambda DN \rceil + N - 3) h\left(\frac{N-2}{\lceil \lambda DN \rceil + N - 3}\right) / \log 2}}{\lceil \lambda DN \rceil + N - 2} \quad (26)$$

where $h(p) = -p\log p - (1-p)\log(1-p)$ and (26) is obtained from the inequality

$$\binom{n}{k} \geq \frac{1}{n+1} 2^{nh(k/n)/\log 2}$$

[10, p. 151]. For each possible delay sum

$$k \in \{0, 1, \ldots, N\lceil \lambda DN \rceil\}$$

let

$$S_k = \left\{ (x_1, \ldots, x_N) \in \mathcal{U}\colon \sum_{j=1}^{N} j x_j = k \right\}.$$

We choose $s = \arg\max_k |S_k|$, so that s is the most common delay sum. Given any $\epsilon > 0$, for N large enough, the number of codewords having delay sum s is lower-bounded by

$$\frac{|\mathcal{U}|}{N\lceil \lambda D N \rceil} \geq 2^{\frac{N}{\log 2}((\lambda D+1)h(\frac{1}{\lambda D+1})-\epsilon)}.$$

One such codeword can be transmitted without error every DN time units, so the rate of the code is at least

$$\frac{(\lambda D+1)h(\frac{1}{\lambda D+1}) - \epsilon}{D} = \frac{H(\mathrm{Geo}_0(\lambda D)) - \epsilon}{D}.$$

Since ϵ is arbitrary, for the MDC jamming channel with delay parameter D and continuous time packet waveforms, we have that

$$C = \overline{V} = \underline{V} = \frac{H(\mathrm{Geo}_0(\lambda D))}{D}.$$

C. MDC, Continuous-Time, Rate-Constrained Fluid Model

The MDC jamming channel for continuous-time, rate-constrained fluid waveforms has infinite capacity for the same reason that the MDC jamming channel for continuous-time fluid waveforms has infinite capacity. That is, $C = \overline{V} = \underline{V} = \infty$.

D. MDC, Discrete-Time Packet Model

An information rate saddle point has not been found for an MDC jammer for discrete-time packet waveforms, but have shown the existence of a saddle point in Section III. We present upper and lower bounds on the saddle-point information rate, $\overline{V}_{DP} = \underline{V}_{DP}$. The best of these upper and lower bounds are within a factor of 3 for $\lambda < 0.7$ and the bounds are reasonably close for $0.7 < \lambda \leq 1$ as shown for $D = 2000$ slots in Fig. 15.

First we describe the jamming strategies that give upper bounds on the maximum saddle-point information rate. Then we describe the input process that gives a lower bound on the saddle-point information rate.

An easy upper bound on the saddle-point information rate comes by using a periodic dump jammer. Using a periodic dump jammer, the output logically takes a value in $0, \ldots, D$ at each dump time, and each packet is delayed by less than D slots. The mean number of packets per dump time is λD, and the distribution on $0, \ldots, D$ with mean λD having the greatest entropy is a geometric distribution, denoted $\mathrm{Geo}_0^D(\lambda D)$, where if Z is distributed according to a $\mathrm{Geo}_0^D(\lambda D)$ random variable, then

$$P(Z = z) = \eta^z / \sum_{i=0}^{D} \eta^i, \qquad \text{for } z \in \{0, \ldots, D\}$$

and η is chosen so that $E[Z] = \lambda D$. Therefore, an upper bound on the entropy rate of the output of an MDC periodic dump jammer and hence an upper bound on the saddle-point information rate is $H(\mathrm{Geo}_0^D(\lambda D))/D$. The bound is plotted in Fig. 15 for $D = 2000$ and labeled *Periodic dump jammer with period* D.

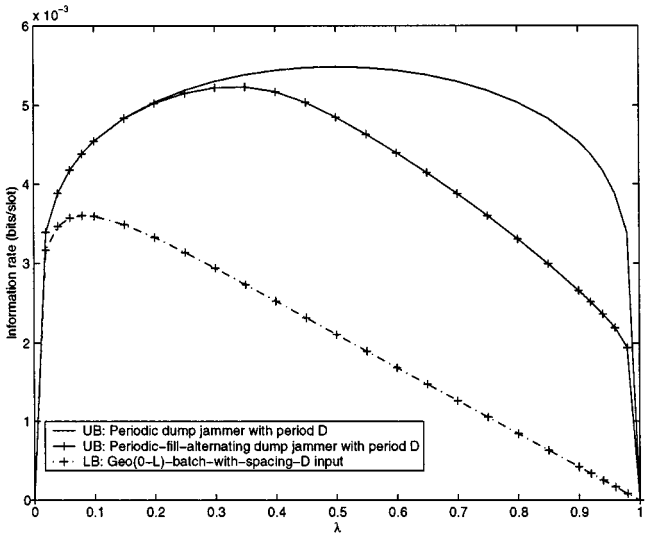

Fig. 15. Information rate for MDC jammer with $D = 2000$, for discrete-time packet waveforms and continuous-time, rate-constrained fluid waveforms.

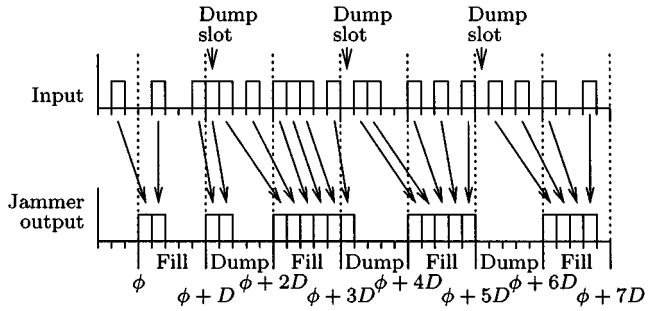

Fig. 16. Fill-alternating periodic dump jammer.

A tighter bound is provided by considering the following jammer. A *fill-alternating periodic dump jammer* with period D for discrete-time packet waveforms is the following variation of a periodic dump jammer for discrete-time packets. For $k \geq 1$, the kth output window is the interval of D slots $[\phi+(k-1)D, \phi+kD)$ (i.e., slots $\phi+(k-1)D$ through $\phi+kD-1$) where the phase ϕ is uniformly distributed on $(0, D]$. In an odd-numbered output window, the jammer transmits a packet in the first slot of the window (if it has one to transmit) and it keeps transmitting packets in each successive slot until either it has no packet to transmit or until the end of the window is reached. Packets arriving after the first slot may still be transmitted, although if the jammer is idle in a slot of the window due to a lack of packets, then it remains idle for the rest of the window. The odd-numbered windows are called *fill* windows since the jammer is trying harder to fill those windows with packets. In an even numbered window the jammer acts the same way as a periodic dump jammer: during the window it transmits only packets that arrive by the first slot of the window. The even-numbered windows are called *dump* windows. An example of a fill-alternating periodic dump jammer is shown in Fig. 16. In effect, a fill window can "steal" some packets from the subsequent dump window.

The maximum entropy rate for the output of a fill-alternating periodic dump jammer is an upper bound on the maximum information rate for the jammer and a rate λ input process. Let Y_k be the number of packets transmitted by the jammer during

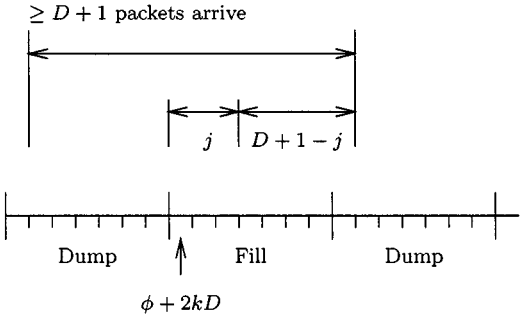

Fig. 17. Fill-alternating periodic dump jammer.

the kth window, $[\phi + (k-1)D, \phi + kD)$, for all $k \geq 1$. Thus, Y_{2k+1} is the number of packets transmitted by the jammer during the fill window $[\phi + 2kD, \phi + (2k+1)D - 1)$, and Y_{2k+2} is the number of packets transmitted during the subsequent dump window. We now prove that with probability one, $(Y_{2k+1}, Y_{2k+2}) \in \mathcal{S}$, where

$$\mathcal{S} = \{(y_1, y_2) \in \{0, 1, \ldots, D\}^2 \colon y_1 + y_2 \leq D \text{ or } y_1 = D\}.$$

Suppose that $Y_{2k+1} + Y_{2k+2} \geq D + 1$. It must be shown that $Y_{2k+1} = D$. At least $D+1$ packets arrive during the interval of $2D$ slots $[\phi + (2k-1)D + 1, \ldots, (2k+1)D]$. Even if these packets arrive as late as possible, for $0 \leq j < D$ at least j of these packets arrive by the jth slot of the $2k+1$th window, as shown in Fig. 17, so that $Y_{2k+1} = D$, as was to be proved.

The maximum entropy distribution on the set $\mathcal{S}$ of possible values for (Y_{2k+1}, Y_{2k+2}) such that $E[Y_{2k+1} + Y_{2k+2}] = 2D\lambda$ is a geometric distribution truncated to $\mathcal{S}$ such that

$$P(Y_{2k+1} = i, Y_{2k+2} = j) = \eta^{i+j} / \sum_{(k,l) \in \mathcal{S}} \eta^{k+l}, \quad \text{for } (i, j) \in \mathcal{S}$$

where η is chosen so that $E[Y_{2k+1} + Y_{2k+2}] = 2D\lambda$. The resulting bound on the saddle-point information rate for an MDC jammer with $D = 2000$ slots is shown in Fig. 15 and is labeled *fill-alternating periodic dump jammer*. Since the fill-alternating periodic dump jammer scheme only deviates from the periodic dump jammer scheme when $Y_{2k+1} + Y_{2k+2}$ is relatively large, the performance of the fill-alternating periodic dump jammer is noticeably better for arrival rates of about $1/2$ and larger.

A lower bound on the saddle-point information rate is given by a $Geo_0^L(\lambda(L+D))$-batch-with-spacing-$(L+D)$ input. The batch sizes for this input range from 0 to L and are chosen independently according to $\mathrm{Geo}_0^L(\lambda(L+D))$ random variables such that for $\lambda < 1$, L is an integer satisfying $L \geq \frac{\lambda D}{1-\lambda}$ (if $\lambda = 1$, no information can be conveyed through timing). Batches are transmitted one packet per slot in consecutive slots starting at the beginning of the $(L+D)$ interval so that there are at least D idle slots between batches. Since the jammer cannot delay packets by D or more time units, the batch sizes can be reproduced exactly at the output, and there is one batch per $L+D$ time units so that the information rate for this input process and any MDC jammer W is at least $\frac{H(\mathrm{Geo}_0^L((L+D)\lambda))}{L+D}$. Since this is the information rate for a particular input in $\mathcal{X}_T$ and any MDC jammer, it is a lower bound on the saddle-point information rate. Fig. 15 illustrates this bound for $D = 2000$ when $L \geq \frac{\lambda D}{1-\lambda}$ is chosen to maximize the bound for each λ and the bound is labeled $Geo(0-L)$*-batch-with-spacing-*D *input*.

It should be noted that the lower bound on the saddle-point information rate obtained from the $\mathrm{Geo}_0^L(\lambda(L+D))$-batch-with-spacing-$(L+D)$ and the upper bound on the saddle-point information rate obtained from the fill-alternating periodic dump jammer have roughly the same general shape and that they differ by a factor of 3 for $\lambda < 0.7$. Note that the lower bound for $\lambda = 0.1$ is greater than the upper bound for $\lambda = 0.9$, implying that the saddle-point information rate is not symmetric around $\lambda = 1/2$.

V. MBC Jammers

In this section, we consider an MBC jammer with parameter B that can carry over at most B units of traffic in its buffer, with the unit of traffic being a packet for the packet models.

A. MBC, Continuous-Time Fluid Model

The capacity for MBC jammers for continuous-time fluid waveforms is infinite. The coder transmits $2B+1$ units of fluid at some time U chosen on $(0, 1]$. Since an MBC jammer can hold no more than B units of fluid, the jammer must immediately release at least $B+1$ units of fluid at time U. Thus, the decoder learns the real-valued U, so an infinite amount of information can be conveyed. Therefore, $C = \overline{V}_{CF} = \underline{V}_{CF} = \infty$ for the MBC jamming channel for continuous-time fluid waveforms.

B. MBC, Continuous-Time Packet Model

Using identical arguments to those in the preceding subsection, $C = \overline{V}_{CP} = \underline{V}_{CP} = \infty$ for the MBC jamming channel for continuous-time packet waveforms.

C. MBC, Continuous-Time, Rate-Constrained Fluid Model

An upper bound on $\overline{V}_{RF}$ for an MBC jammer is obtained with the periodic quantized dump jammer with period and quanta size both equal $B/2$. If the amount of fluid in the buffer at a dump time is less than $B/2$, then the jammer holds the fluid until the next dump time, and at most $B/2$ additional units of fluid can arrive by the next dump time since the input is limited to rate 1. If instead the amount of fluid in the buffer is greater than or equal to $B/2$ at a given dump time, then the jammer will release fluid at rate 1 until the next dump time. Since the input is limited to rate 1 during this interval, the fluid level cannot increase. Since the buffer is initially empty, the fluid level in the jammer's buffer will never exceed B units and the jamming strategy is MBC for continuous-time, rate-constrained inputs.

The jammer's output is logically one of two symbols every $B/2$ time units, regardless of the input process. The information rate for any rate λ input process and this MBC jamming strategy is upper-bounded by the maximum entropy rate of the jammer's output, $2h(\lambda)/B$. Hence, an upper bound on $\overline{V}_{RF}$ is given by $2h(\lambda)/B$.

We show in the following subsection that capacity for the MBC jamming channel for discrete-time packet waveforms is lower-bounded by $\frac{h(\lambda)}{B+1}$. We can achieve the same rate for continuous-time, rate-constrained fluid waveforms by treating the

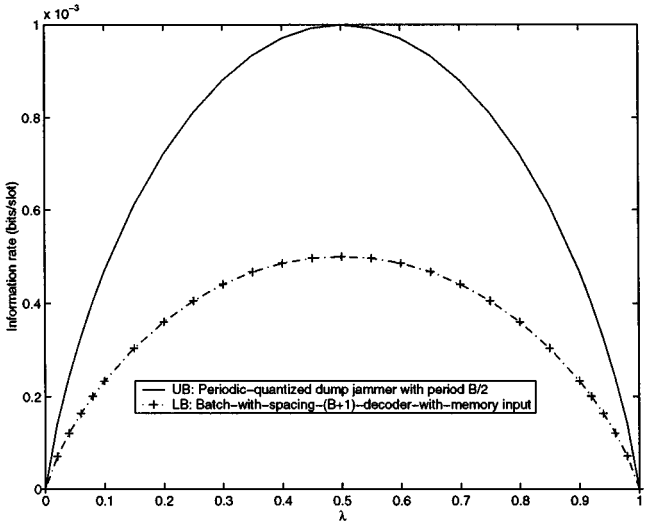

Fig. 18. Bounds on information rate for MBC jammers, $B = 2000$.

fluid as packets and placing a trigger packetizer and slotter at the jammer output (start the rate-constrained waveform at time 1). Thus, $\frac{2h(\lambda)}{B} \geq C \geq \frac{h(\lambda)}{B+1}$ for the MBC jamming channel for continuous time, rate-constrained fluid waveforms.

D. MBC, Discrete-Time Packet Model

Upper bounds on $\overline{V}_{DP}$ and lower bounds on $\underline{V}_{DP}$ for MBC jamming channels are presented in this subsection and are illustrated in Fig. 18. The upper and lower bounds that we have found are within a factor of 2 of each other for all values of λ.

Let $\mathcal{W}$ denote the set of MBC jammers for discrete-time packet waveforms and assume that the jammer buffer is always initially empty. Let $\mathcal{X}_T$ be the set of rate λ discrete-time packet waveforms.

An upper bound on $\overline{V}_{DP}$ for an MBC jammer is obtained with the periodic quantized dump jammer with period and number of packets per quanta both equal to $\lfloor \frac{B}{2} \rfloor + 1$. Let

$$\phi_i = \phi + (i-1)\left(\left\lfloor \frac{B}{2} \right\rfloor + 1\right)$$

which is the ith dump slot. If the jammer has $\lfloor \frac{B}{2} \rfloor$ or fewer packets in its buffer at ϕ_i including the arrival in that slot, then the jammer holds all packets until ϕ_{i+1} and at most B packets will be in the jammer just before ϕ_{i+1}. If, instead, the jammer has more than $\lfloor \frac{B}{2} \rfloor$ packets in its buffer at ϕ_i, then the jammer will release one packet in each slot during the period from ϕ_i to just before ϕ_{i+1} and the total buffer size will not increase before ϕ_{i+1}. Hence, this jamming strategy will never hold over more than B packets and it is an MBC jamming strategy. The jammer output is either 0 or $\lfloor \frac{B}{2} \rfloor + 1$ packets in a $(\lfloor \frac{B}{2} \rfloor + 1)$ interval, so the output can be considered a discrete binary process taking values 0 or $\lfloor \frac{B}{2} \rfloor + 1$ at each ϕ_i. The entropy rate for the output will be maximized if the outputs at the dump slots are independent. The mean number that should depart at each ϕ_i is $\lambda(\lfloor \frac{B}{2} \rfloor + 1)$ since the arrival rate is λ. Hence, the output entropy rate and therefore the min-max information rate for the channel model, is upper-bounded as $\overline{V}_{DP} \leq \frac{h(\lambda)}{\lfloor \frac{B}{2} \rfloor + 1} \leq \frac{2h(\lambda)}{B+1}$. The bound is illustrated in Fig. 18 for $B = 2000$ packets and is labeled *periodic quantized dump jammer*.

A simple lower bound on $\underline{V}_{DP}$ is obtained with the following batch-with-spacing-$(B+1)$ process. Time is divided into super-slots consisting of $(B+1)$ consecutive slots. In each super-slot, the encoder with probability $1-\lambda$ transmits a binary 0 by not sending any packets, and otherwise it transmits a binary 1 by filling the super-slot with $B+1$ packet transmissions. We will show that the input is determined by the jammer output. The decoder makes hard decisions at the end of each super-slot. We define outstanding packets as any packets sent by the encoder, but not yet received by the decoder. We claim that at each decision time, the decoder has sufficient information to make the correct hard decision and to determine the number of outstanding packets. The proof is by induction. If the first transmitted bit is a zero, then zero packets are output in the first super-slot. If the first transmitted bit is a one, then at least one packet is transmitted in the first super-slot. Thus, the claim is true for the first decision time $B+1$. Suppose the claim is true for the kth decision time. In particular, the decoder knows the number j of outstanding packets at the kth decision time. If the encoder sends a binary 0 in the $k+1$th super-slot, then the jammer must output j or fewer packets in the $k+1$th super-slot. If the encoder sends a binary 1 in the $k+1$th super-slot, then the jammer must output at least $j+1$ packets in the $k+1$th super-slot. Thus, at the end of the $k+1$th super-slot, the decoder will know the information bit sent during the $k+1$th super-slot. Using that knowledge, together with knowing j and the number of packets input during the $k+1$th super-slot, the decoder can calculate the number of outstanding packets. Thus, the proof of the claim by induction is concluded. Since the decoder receives one error-free binary symbol every $B+1$ slots and the arrival rate is λ, we have that both $\underline{V}_{DP}$ and C are lower bounded by $\frac{h(\lambda)}{B+1}$. This lower bound differs from the upper bound by a factor of 2 and is illustrated in Fig. 18 with the label *batch-with-spacing-$(B+1)$-decoder-with-memory input* for $B = 2000$.

To close this subsection, we explore another philosophy for jammers. The idea is for the jammer to try to make the output process as random as possible. For example, the jammer might like to output a stream of independent, mean λ Bernoulli variables, regardless of the input. Of course, this is not possible since the buffer can become full or empty. A reasonable fix would be for the jammer to adjust the output probability as a function of the buffer size. For example, consider a jammer that in any slot i outputs a packet with probability $\frac{k}{B+1}$, where k is the number of packets in the buffer at the beginning of the slot, including the new arrival, if any. However, this is apparently not a good jammer strategy, at least for B large. Indeed, this jammer is equivalent to the generalized billiard ball channel of Berger [5].

In Berger's model, the coder adds a red or white ball to a billiard table at the beginning of each slot, and the channel output is a ball selected from the table, with all balls having equal probability. The equivalence can be seen by mapping slots with packets to red balls, and idle slots to white balls. It can be shown [6] that for large B the capacity of the channel is at least as large as a constant times $B^{-\frac{2}{3}}$. For large B, this far exceeds the capacity upper bound $\frac{2h(\lambda)}{B+1}$ for the periodic quantized dump jammer described above. Roughly speaking, the generalized billiard ball channel can allow the coder to reliably convey infor-

mation by causing small fluctuations in buffer size over short time intervals.

VI. ADC JAMMERS

The class of ADC jammers is considered in this section.

A. ADC, Continuous-Time Fluid Model

By Theorem III.2, $\overline{V}_{CF}$ and $\underline{V}_{CF}$ are each given by a constant divided by D. The upper bound on $\overline{V}_{CF}$ and the lower bound on $\underline{V}_{CF}$ that are obtained for the strategies we present differ by a factor of about 8.

Choose ϵ such that $0 \leq \epsilon < \frac{\lambda - \epsilon}{\lambda}$ and consider the class of rate $\lambda \pm \epsilon$ input processes $\mathcal{X}_{T,\epsilon}$. An upper bound on $\overline{V}_{CF}$ for an ADC jammer is obtained with a periodic quantized dump jammer with period LD, quantized to $\alpha\lambda D$, where $L = 1 - \frac{\alpha\lambda}{\lambda - \epsilon}$, and $0 < \alpha < 1$. The constant α is chosen to obtain the smallest upper bound. Note that *the jammer must know* λ for this strategy. We will first show that a periodic quantized dump jammer with period LD, quantized to $\alpha\lambda D$, is an ADC jammer for continuous-time fluid waveforms.

For a drop of fluid, the time it spends in the system can be broken into two parts: 1) an initial delay equal to the time from its arrival until the first dump time after its arrival, and 2) a carryover delay equal to the time that the fluid waits from the first dump time until the last dump time that the fluid is in the system. Since the spacing between dump times is LD, the largest amount of initial delay experienced by a drop of fluid is LD. The average amount of type 2 delay experienced by a drop of fluid is at most $\frac{\alpha D\lambda}{\lambda - \epsilon}$ by Little's law, since less than $\alpha\lambda D$ units of carryover fluid is in the buffer at any time. Therefore, the mean delay experienced by fluid is at most $LD + \frac{\alpha D\lambda}{\lambda - \epsilon} = D$ time units.

The maximum information rate for a rate λ input process and the periodic quantized dump jammer is upper-bounded by the maximum jammer output entropy rate. The jammer has batch outputs every LD time units where the batch size is 0 or a positive integer multiple of $\alpha\lambda D$. The average batch size must be at most $L(\lambda+\epsilon)D$ units of fluid since the batches occur every LD time units and the packet input rate is at most $\lambda + \epsilon$. Arguing as in the proof of Theorem IV.1 yields that the maximum output entropy rate for a periodic quantized dump jammer with period LD, quantized to $\alpha\lambda D$ is given by $H(\text{Geo}_0(\frac{L(\lambda+\epsilon)}{\alpha\lambda}))/(LD)$. Minimizing this expression with respect to α and taking the limit as ϵ tends to 0 yields an upper bound for $\overline{V}_{CF}$. The optimizing α tends to $1/2$, and the bound is

$$\overline{V}_{CF} \leq \frac{4 \text{ bits}}{D}. \tag{27}$$

Theorem III.5 guarantees that (27) is also an upper bound on $\overline{V}_{CP}$ for ADC jamming channels.

Many of our best input processes attempt to force the output to allow good prediction of the input at the expense of using input processes that are closer to deterministic, especially with respect to timing. Our best input process for an ADC jammer for continuous-time fluid waveforms, however, takes the opposite approach and selects fluid departure times at random. We first state the following theorem which is proved in the Appendix.

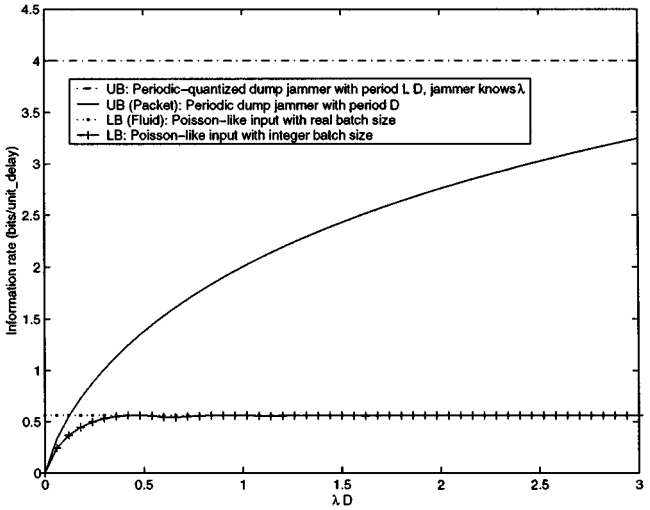

Fig. 19. Normalized information rate (bits/unit delay) for an ADC jammer, with continuous-time packet waveforms or continuous-time fluid waveforms.

Theorem VI.1: Assume continuous-time packet waveforms or continuous-time fluid waveforms. Given k, $T > 0$, let X be a Poisson-like input with batch size k with packet arrival rate $\lfloor \lambda T/k \rfloor k/T$, and let $\mathcal{W}_T$ be a class of ADC jamming channels on $[0, T]$. Then

$$\liminf_{T\to\infty} \inf_{W \in \mathcal{W}_T} \overline{I}_T(X, W) \geq -\frac{\lambda}{k} \log\left(\frac{\lambda\eta}{k} e^{\frac{D}{\eta}-1}\left(1 - \frac{1}{2} e^{-\frac{k}{\lambda\eta}}\right)\right) \tag{28}$$

where $\eta > 0$ can be chosen to maximize the bound. The batch size k may be chosen to be any real-valued, positive number in the continuous-time fluid model. Taking $\alpha = \frac{\lambda D}{k}$ and $\beta = \frac{D}{\eta}$

$$\underline{V}_{CF} \geq \sup_{\alpha,\, \beta \in \Re^+} -\frac{\alpha}{D} \log\left(\frac{\alpha}{\beta} e^{\beta-1}\left(1 - \frac{1}{2} e^{-\frac{\beta}{\alpha}}\right)\right) \tag{29}$$

$$\approx \frac{0.5615 \text{ bits}}{D} \tag{30}$$

which holds for all $\lambda > 0$ with optimum batch size $k \approx 2.15\lambda D$. The bound of (30) is illustrated in Fig. 19 of Section VI-B and is labeled *Poisson-like input with real batch size*.

B. ADC, Continuous-Time Packet Model

An MDC jammer for maximum delay D is also an ADC jammer for average delay D. In particular, a periodic dump jammer with period D satisfies the MDC constraint. Therefore, a slight modification of the proof of Theorem IV.1 to account for the ADC input constraints, implies that

$$\overline{V}_{CP} \leq \frac{H(\text{Geo}_0(D\lambda))}{D}.$$

Another upper bound was already mentioned in Section VI-A. A jammer that quantizes output batch sizes (using knowledge of the input rate λ) yields $\overline{V}_{CP} \leq \overline{V}_{CF} \leq \frac{4}{D}$. The two upper bounds are pictured in Fig. 19. The periodic jump jammer gives the tighter bound for small λD, whereas for large λD, when batches tend to be large, a jammer that quantizes output batch sizes is better.

For a lower bound on $\underline{V}_{CP}$, we can use the Poisson-like-input with batch size k from Section VI-A with batch size k an integer. More precisely, select an integer $k \geq 1$ to minimize the right-hand side of (28). This lower bound is illustrated in Fig. 19 and is labeled *Poisson-like input with integer batch size*.

C. ADC, Continuous-Time, Rate-Constrained Fluid Model

An upper bound on $\overline{V}_{RF}$ for an ADC jammer is obtained with a periodic quantized dump jammer with period LD, quantized to $\alpha\lambda D$ similar to that of Section VI-A, where L and α are carefully chosen constants depending on λ and D such that $L/(\alpha\lambda)$ is an integer $L = \frac{2}{3}(1 - \frac{\alpha\lambda}{\lambda-\epsilon})$, and $0 < \alpha < \frac{\lambda-\epsilon}{\lambda}$. For this strategy, we assume that the jammer *knows the nominal rate λ of the input process*. We will first show that such a jammer is ADC and then give an upper bound on the maximum information rate for this jammer.

The average delay experienced by fluid in this jammer can be obtained in the following way. For a drop of fluid, the time it spends in the system can be broken into three parts: 1) the time from its arrival until the first dump time, 2) the time from the first dump time until the last dump time the fluid is in the system (this is the carryover period for the drop of fluid), and 3) the time from the last dump time until the departure of the drop of fluid. The type 1 delay is at most LD time units since dump times have spacing LD. The average type 2 delay is at most $\frac{\alpha D\lambda}{\lambda-\epsilon}$ by Little's law, since, the amount of fluid carried between any two dump times is less than $\alpha\lambda D$, and the arrival rate is at least $\lambda - \epsilon$. Finally, the type 3 delay for the drop of fluid is on average at most $\frac{LD}{2}$ since the drop may depart at any time in the LD interval. Thus, the mean delay is at most

$$LD + LD/2 + \alpha D\lambda/(\lambda - \epsilon) = D$$

so the jamming scheme is ADC.

The mean number of quanta of size $\alpha\lambda D$ transmitted at each dump time is at most

$$\mu = \frac{L(\lambda+\epsilon)D}{\alpha\lambda D} = L(\lambda+\epsilon)/(\alpha\lambda)$$

and the largest number of quanta that can be transmitted is at most

$$\frac{LD}{\alpha\lambda D} = \frac{L}{\alpha\lambda}.$$

Arguing as in the proof of Theorem IV.1 yields

$$\overline{V}_{RF} \leq \frac{H\left(\mathrm{Geo}_0^{L/(\alpha\lambda)}\left(\frac{L}{\alpha}\right)\right)}{LD}. \tag{31}$$

By Theorem III.7, $\overline{V}_{RF} \geq \overline{V}_{DP}$ so the bound in (31) also holds for discrete-time packet waveforms. The bound in (31) is illustrated for $D = 8$ in Fig. 21 of Section VI-D for optimal L and α and is labeled *periodic quantized dump jammer with period LD, jammer knows λ*.

To obtain a lower bound on $\underline{V}_{RF}$, we use a Bernoulli-like-input with batch size k. The following theorem is proved in the Appendix.

Theorem VI.2: Assume continuous-time, rate-constrained fluid waveforms or discrete-time packet waveforms. Given $k > 0$ and $T > 0$ with T/k an integer, let X be a Bernoulli-like-input with batch size k over $[0, T]$ with packet arrival rate $\lfloor \lambda T/k \rfloor k/T$ and let $\mathcal{W}_T$ be the class of ADC jamming channels. Then for each $\eta > 0$,

$$\liminf_{T\to\infty} \inf_{W\in\mathcal{W}_T} \overline{I}_T(X, W) \geq \frac{h(\lambda)}{k} - \frac{\lambda}{k}\log\left(\eta^{-1}(1-\eta)^{-\frac{D}{k}}\,\phi\left(\frac{1}{\lambda}, \eta\right)\right) \tag{32}$$

where $\phi(u, \eta) = 1 - \frac{(1-\eta)^u}{2-\eta}$.

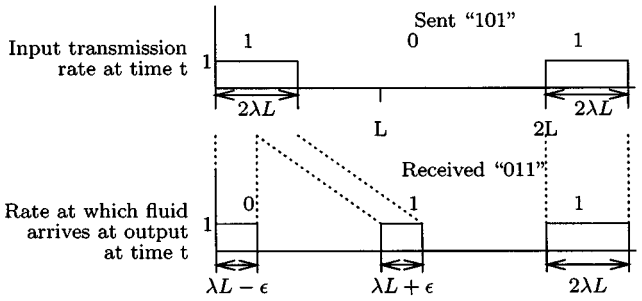

Fig. 20. Symbol errors caused by delay.

For continuous-time, rate-constrained fluid waveforms, we may choose the batch size k to be real-valued. Maximizing (32) with respect to k and η, we obtain a lower bound on $\underline{V}_{RF}$ which is illustrated in Fig. 21 of Section VI-D and is labeled *Bernoulli-like input with real-valued batch size*.

Next, we present a coding scheme that gives a lower bound on capacity for an ADC jamming channel for continuous-time, rate-constrained fluid waveforms.

For this coding scheme, we take a time interval of length L to represent a single channel use, and assume $\lambda < 1/2$. To send a binary 1, we transmit $2\lambda L$ units of fluid in the L interval, and to transmit a binary 0 we transmit zero units of fluid in an L interval. The decoder decides that a 1 was transmitted if it sees λL units of fluid or more in an L interval, and a 0 was transmitted if it sees less than λL units of fluid. Note that we are assuming the coder and decoder have access to a common clock.

Given a large even integer N, the set of codewords is a subset of the length N binary sequences with normalized weight (fraction of 1's) equal to $1/2$. The code waveforms have duration NL. The decoder makes symbol-by-symbol decisions, and then the resulting received binary sequence is decoded to the nearest codeword. As illustrated in Fig. 20, if an amount of fluid $\lambda L + \epsilon$ is delayed by $L - \lambda L - \epsilon$ time units, then two bit errors can result. The product of the amount of fluid moved and the delay is $(\lambda L+\epsilon)(L-\lambda L-\epsilon)$. No output with smaller fluid-delay product can cause an error. Moreover, any number of errors would require a total fluid-delay product of at least $\lambda L^2(1-\lambda)/2$ per error. On the other hand, the ADC constraint ensures that there is at most λLND fluid delay per codeword, so the number of errors per codeword is at most $\frac{2D}{L(1-\lambda)}N$.

By choosing the set of codewords so that the minimum distance of the code is at least twice the maximum number of errors per codeword, the decoder can make error-free minimum distance decisions. The Gilbert–Varshamov bound guarantees that for N large enough, a binary code with relative weight $1/2$, rate $\tilde{R}$, and minimum distance pN exists, provided $\tilde{R} <$

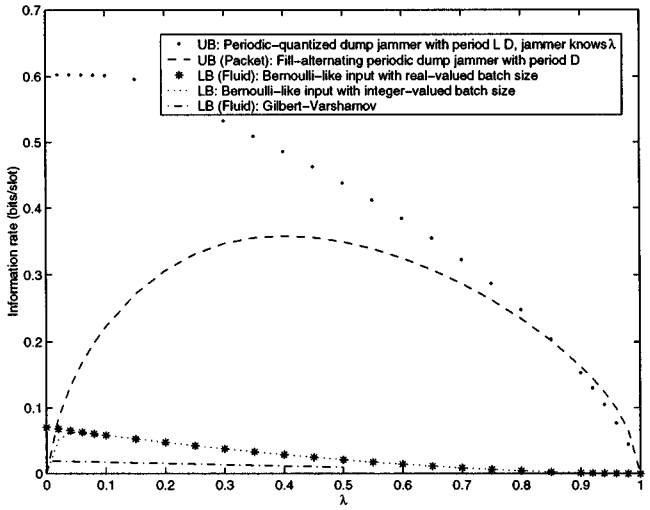

Fig. 21. Information rate for and ADC jammer with $D = 8$, for discrete-time packet waveforms or continuous-time, rate-constrained fluid waveforms.

$\log(2) - h(p)$. Then taking $p = \frac{4D}{L(1-\lambda)}$, for $\delta > 0$ and N large enough, we choose a collection of M such codewords from the collection of binary sequences, where

$$\frac{1}{N}\log M \geq \log(2) - h\left(\frac{4D}{L(1-\lambda)}\right) - \delta.$$

Since M symbols can be transmitted without error every LN time units

$$C \geq \frac{\log 2 - h\left(\frac{4D}{L(1-\lambda)}\right)}{L} \tag{33}$$

for ADC jammers. The time length for a symbol L can be chosen to maximize the bound in (33).

A slight improvement to this bound on capacity can be made by choosing codewords with relative weight $0 < w < 1$ and transmitting a binary 1 using $\frac{\lambda L}{w}$ units of fluid in an L interval with probability w and a 0 with no fluid in an L interval with probability $1 - w$. We obtain a lower bound depending on L, w, λ, and D which makes use of a version of the Gilbert–Varshamov bound for weighted codewords that guarantees, for N large enough, a binary code with relative weight w, rate $\tilde{R}$, and minimum distance pN exists, provided

$$\tilde{R} < h(w) - wh\left(\frac{p}{2w}\right) - (1-w)h\left(\frac{p}{2(1-w)}\right).$$

The bound can then be optimized over L and w and is nonzero for $0 < \lambda < 1$. The weight $1/2$ version of the bound is illustrated in Fig. 21 of Section VI-D for $D = 8$ slots and is labeled *Gilbert–Varshamov*.

D. ADC, Discrete-Time Packet Model

In this subsection we give the jamming strategies and input processes for ADC jammers for the discrete-time packet waveforms. The bounds for the discrete-time packet waveforms and continuous-time, rate-constrained fluid waveforms are illustrated in Fig. 21 for $D = 8$ slots.

An MDC jammer is an ADC jammer. In particular, the upper bound on $\overline{V}_{DP}$ for MDC jammers obtained in Section IV-D by consideration of a fill-alternating periodic dump jammer in Section IV-D is also an upper bound on $\overline{V}_{DP}$ for ADC jammers.

In addition, from Theorem III.7, any upper bound on $\overline{V}_{RF}$ for an ADC jammer is an upper bound on $\overline{V}_{DP}$ for an ADC jammer. Thus, for the case when *the jammer knows the input packet transmission rate* λ, we may use the periodic quantized dump jammer with period LD, quantized to $\alpha\lambda D$ jamming strategy of Section VI-C by preceding that jammer with a fluidizer and following the jammer with a trigger packetizer and slotter as in Fig. 7.

For a lower bound on $\underline{V}_{DP}$, we can use the Bernoulli-like input with batch size k, discussed in Section VI-C, where k is taken to be an integer. Maximizing the bound of (32) with respect to k and η where k is an integer, we obtain the corresponding lower bound which is illustrated in Fig. 21 for $D = 8$ slots labeled *Bernoulli-like input with integer-valued batch size*.

VII. Discussion

In this section, we summarize the results. In particular, we give rules of thumb for jammers, rules of thumb for inputs, and numerical examples. In addition, we discuss the use of the ideas from this paper in developing timing channel coding schemes for use on the Internet.

A. Summary of Results

Table I summarizes all of the jammed timing channel models that are considered in this paper. In the table, the asymptotic behavior for the saddle-point information rate (if a saddle point exists) and for capacity is described in terms of standard Θ-notation where $\Theta(g(n))$ represents the set of functions

$$\{f(n)\text{: there exists } c_1 > 0,\ c_2 > 0,\ n_0 > 0 \text{ such that } c_1 g(n) \leq f(n) \leq c_2 g(n) \text{ for all } n \geq n_0\}.$$

Table II summarizes the good jamming strategies we have found for each of the jamming channel models and Table III summarizes the good input strategies we have found for each of the jamming channel models.

In every model considered, the best jammers that we have found are those that have some sort of quantized batch departures at regular intervals. By only allowing outputs at regular intervals, the jammers eliminate much of the output uncertainty derived from timing. By requiring quantized output, the jammers eliminate much of the output uncertainty derived from variations in the output intensity.

Since jammers of this type are generally deterministic, the information rate between a particular coding scheme input and the output of the jammer is given by the output entropy rate. Such a jammer would want to make the output entropy as small as possible, and thus, quantized batches at regular intervals make sense.

For many of the waveforms and jammer constraints considered, the best inputs and coding strategies that we have found are those that have batch arrivals at regular intervals. If information were encoded in the input on a finer time scale it would be easily concealed by a jammer anyway. Another benefit of these batch strategies is that by spacing the batches at regular intervals with large enough spacing between batches, error-free

TABLE I
SUMMARY OF CAPACITY AND INFORMATION RATE RESULTS

Jammers \ Waveforms	Continuous time fluid waveforms	Continuous time packet waveforms	Continuous time, rate-constrained fluid waveforms	Discrete time packet waveforms
Maximum-delay-constrained	Subsection IV.A. $C = \overline{V} = \underline{V} = \infty$.	Subsection IV.B. *Saddle point.* $C = \overline{V} = \underline{V}$. C is $\Theta(\log D/D)$.	Subsection IV.C. $C = \overline{V} = \underline{V} = \infty$.	Subsection IV.D. $C \leq \overline{V} = \underline{V}$. $\overline{V}$ and $\underline{V}$ are $\Theta(\log D/D)$.
Maximum-buffer-constrained	Subsection V.A. $C = \overline{V} = \underline{V} = \infty$.	Subsection V.B. $C = \overline{V} = \underline{V} = \infty$.	Subsection V.C. $C \leq \overline{V}$. C, $\overline{V}$, and $\underline{V}$ are $\Theta(1/B)$.	Subsection V.D. $C \leq \overline{V}$. C, $\overline{V}$, and $\underline{V}$ are $\Theta(1/B)$.
Average-delay-constrained (jammer knows λ)	Subsection VI.A. $C \leq \overline{V}$. $\overline{V}$ and $\underline{V}$ are $\Theta(1/D)$.	Subsection VI.B. $C \leq \overline{V}$. $\overline{V}$ and $\underline{V}$ are $\Theta(1/D)$.	Subsection VI.C. $C \leq \overline{V}$. $\overline{V}$ and $\underline{V}$ are $\Theta(1/D)$.	Subsection VI.D. $C \leq \overline{V}$. $\overline{V}$ and $\underline{V}$ are $\Theta(1/D)$.

TABLE II
SUMMARY OF JAMMER USAGE

Jammers \ Waveforms	Continuous time fluid waveforms	Continuous time packet waveforms	Continuous time, rate-constrained fluid waveforms	Discrete time packet waveforms
Maximum-delay-constrained	Irrelevant – infinite capacity.	Periodic dump jammer with period D.	Irrelevant – infinite capacity.	Fill-alternating periodic dump jammer with period D.
Maximum-buffer-constrained	Irrelevant – infinite capacity.	Irrelevant – infinite capacity.	Periodic quantized dump jammer with period $B/2$ quantized to $B/2$.	Periodic quantized dump jammer with period $\lfloor \frac{B}{2} \rfloor + 1$, quantized to $\lfloor \frac{B}{2} \rfloor + 1$.
Average-delay-constrained (λ known)	Periodic quantized dump jammer with period LD, quantized to $\alpha\lambda D$, $L = (1-\alpha)$, $0 < \alpha < 1$.	Periodic quantized dump jammer with period LD, quantized to $\alpha\lambda D$, $L = (1-\alpha)$, $0 < \alpha < 1$. (Followed by trigger packetizer.); Periodic dump jammer with period D for small λD.	Periodic quantized dump jammer with period LD, quantized to $\alpha\lambda D$, $L/(\alpha\lambda)$ an integer, $L = (2/3)(1-\alpha)$, and $0 < \alpha < 1$.	Periodic quantized dump jammer with period LD, quantized to $\alpha\lambda D$, $L/(\alpha\lambda)$ an integer, $L = (2/3)(1-\alpha)$, and $0 < \alpha < 1$. (Preceded by fluidizer, followed by trigger packetizer and slotter.)
Average-delay-constrained (λ unknown)	Unknown.	Periodic dump jammer with period D.	Unknown.	Fill-alternating periodic dump jammer with period D.

decoding can often occur in the case of MDC or MBC jammers. Thus, the information rate between the input and output is equal to the entropy rate of the input and we can select batch strategies with maximum entropy.

For ADC jamming channels, the best strategies that we have found use quantized batch arrivals, but the batch times are chosen randomly rather than occurring at deterministic, regularly spaced times. However, we cannot say with much confidence that this type of input is good for ADC jamming channels (especially for discrete packet waveforms) since the gap between the upper and lower bounds is so large.

B. Numerical Example

Suppose a transmitter tries to covertly use packet timing over a 100-Mb/s link with 10 000 bit packets. For simplicity, assume no other traffic is on the link, and that discrete-time packet waveforms are used with time divided into intervals of length 10^{-4} s each. In addition, assume that any delay introduced beyond that introduced by a jammer is known to both the transmitter and receiver. Take the long-run transmission rate to be $\lambda = 1/3$, corresponding to a base packet channel bit rate of 33.3 Mb/s. The capacity of this timing channel *without a jammer* is given by $h(1/3)/10^{-4} \approx 9200$ b/s, corresponding to a transmitter that transmits a packet in each slot with probability $1/3$, independently of other slots.

Sections IV-B–VI-B provide upper bounds on $\overline{V}$ and lower bounds on $\underline{V}$ for various constraints on the jammer. For rate $1/3$ input processes and MDC jammers with maximum delays 2 ms, 20 ms, 200 ms, and 2 s, MBC jammers with maximum buffer sizes 20, 200, 2000, and 20 000, and ADC jammers with average delays 2 ms, 20 ms, 200 ms, and 2 s, we present bounds

TABLE III
SUMMARY OF INPUT USAGE

Jammers \ Waveforms	Continuous time fluid waveforms	Continuous time packet waveforms	Continuous time, rate-constrained fluid waveforms	Discrete time packet waveforms
Maximum-delay-constrained	Batch-with-spacing-D input, real-valued batch size.	$Geo_0(\lambda D)$-batch-with-spacing-D.	Batch-with-spacing-D input, real-valued batch size.	$Geo_0^L(\lambda(L+D))$-batch-with-spacing-$(L+D)$.
Maximum-buffer-constrained	Send $2B+1$ units at once, decoder recovers exact time (real-valued) when it sees $\geq B+1$ units.	Send $2B+1$ packets at once, decoder recovers exact time (real-valued) when it sees $\geq B+1$ packets.	Batch-with-spacing-$(B+1)$ input.	Batch-with-spacing-$(B+1)$ input.
Average-delay-constrained (λ known or unknown)	Poisson-like-input with batch size k with k real-valued.	Poisson-like-input with batch size k with k integer-valued.	Bernoulli-like-input with batch size k with k real-valued.	Bernoulli-like-input with batch size k with k integer-valued.

TABLE IV
DISCRETE-TIME NUMERICAL EXAMPLE (100-Mb/s LINK, 33.3-Mb/s HOST STREAM BIT RATE, 10 000 BITS/PACKET)

Jammers \ Delay (Buffer Size)	D=2 ms (B=20 packets)		D=20 ms (B=200 packets)		D=200 ms (B=2,000 packets)		D=2,000 ms (B=20,000 packets)	
	UB	LB	UB	LB	UB	LB	UB	LB
No Jammer	9183 bps							
Maximum-delay-constrained	1985 bps	1073 bps	358 bps	188 bps	52 bps	28 bps	6.9 bps	3.7 bps
Maximum-buffer-constrained	875 bps	437 bps	91 bps	45 bps	9.2 bps	4.5 bps	0.92 bps	0.45 bps
Average-delay-constrained (λ known)	2066 bps	138 bps	207 bps	13.8 bps	20.7 bps	1.38 bps	2.07 bps	0.138 bps
Average-delay-constrained (λ unknown)	1985 bps	138 bps	358 bps	13.8 bps	52 bps	1.38 bps	6.9 bps	0.138 bps

in Table IV. Note that the delays correspond to 20, 200, 2000, and 20 000 slots since the slots are 10^{-4} s. The results show, for example, that for an MDC jammer with maximum delay of 20 ms, jamming strategies exist for which no coding strategy can exceed 358 b/s through the jammer. Additionally, for rate $1/3$ coders and MDC jammers with maximum delay 20 ms, coding strategies exist for which at least 188 b/s can be transmitted, regardless of the particular choice of MDC jamming strategy. Thus, if saddle-point strategies exist for this example, the saddle-point information rate with such strategies is between 188 and 358 b/s for MDC jammers with maximum delay 20 ms and link utilization of $1/3$.

As can be seen from the table, for relatively small delay or small buffers, the maximum rate at which information can be transmitted using the timing channel in the presence of good jammers is reduced significantly.

C. Conclusions

The primary conclusion of this work is that timing channel jammers which use batching and quantizing schemes with deterministic batch departure times generally perform well for the models considered. One such jammer channel and a batch input process with deterministic batch arrival times were shown to be a saddle point for an MDC jammer for continuous-time packet waveforms. In addition, for all of the other models considered, reasonably tight bounds on the information rate for saddle points have been provided. In many cases, channel capacity has been bounded and relationships among C, $\overline{V}$, and $\underline{V}$ have been illustrated. Table I in Section VII-A summarizes the results.

D. Further Research

We list potential areas of further research below.

- *Tighten bounds for existing models*. We have found a saddle point only for the MDC jammer for continuous-time packet waveforms.
- *Investigate other versions of the jammed timing channel problem*. There are many other versions of the delay channel problem we can investigate. For example, we could consider other input classes such as the (σ, ρ) upper constrained inputs [11].
- *Consider a statistical model for jammers*. Investigate the timing channel for statistical jammers, such as the Pump [29], [20].
- *Investigate other covert channels*. Embedding information in timing is one of many ways that information may be hidden. For packet channels, there are many ways

that information may be hidden such as modulating the packet length, modifying packet headers, or purposely introducing bit errors.

- *Implementations*. A covert timing channel jammer could be implemented for Internet traffic. Such a jammer could be incorporated into a network firewall.

APPENDIX

A. Proof of Theorem VI.1

By construction, the number of batches in the input X is given by $N = \lfloor \lambda T/k \rfloor$. Since batches are transmitted instantaneously and the batch sizes are the same, only the batch transmission times are informative. Denote the arrival times of the N batches of X as $X_1, X_2, \ldots, X_N$. Let Y represent the output of some $W \in \mathcal{W}_T$ when X is the input (first pass the jammer output through a trigger packetizer that groups fluid into k-sized "packets"). Without loss of generality, we assume all N packets are output by time T. Thus, the output Y can be represented by the batch arrival times of the output $(Y_1, \ldots, Y_N)$.

Let

$$S = \{(x_1, \ldots, x_N) \colon 0 < x_1 \le x_2 \le \cdots \le x_N \le T\}.$$

The joint density for $(X_1, \ldots, X_N)$ is given by $\frac{N!}{T^N}$ on S and zero outside S by the input construction. Thus, the entropy of X relative to the Lebesgue measure on $\Re^N$ is given by

$$\begin{aligned} H(X) &= -\int \cdots \int_{S_T} \frac{N!}{T^N} \log \frac{N!}{T^N}\, dx_1 \cdots dx_N \\ &= \log \frac{T^N}{N!}. \end{aligned}$$

Now let $y \in S$. Think of y as a particular output for some jammer in $\mathcal{W}_T$ and some input $X \in S$. Let

$$A_y = \{x \in S \colon W_x = y \text{ for some } W \in \mathcal{W}_T\}.$$

Then A_y represents the set of inputs in S for which y is a valid output under some jammer in $\mathcal{W}_T$. The distribution supported by A_y with the greatest relative entropy is the uniform distribution over the set. Thus, $H(X|Y=y) \le \log |A_y|$ since $\log |A_y|$ is the relative entropy of a uniform distribution on A_y. Therefore, we have that

$$H(X|Y) \le \max_y H(X|Y=y) \le \max_y \log |A_y|.$$

Fix an output y and let $\overline{D} = (\overline{D}_1, \ldots, \overline{D}_N)$ be independent uniform random variables on $[0, T]$. We have that

$$\begin{aligned} |A_y| &= T^N P\left(y - \overline{D} \in A_y\right) \\ &= \int \cdots \int_{[0,T]^N} 1_{\{R\}} T^N f_{\overline{D}}\left(\overline{d}_1, \ldots, \overline{d}_N\right)\, d\overline{d}_1 \cdots d\overline{d}_N \\ &= \int \cdots \int_{[0,T]^N} 1_{\{R\}}\, d\overline{d}_1 \cdots d\overline{d}_N \qquad (34) \end{aligned}$$

where $1_{\{R\}}$ is the indicator function taking value 1 if R is true and 0 otherwise, with

$$R = \left\{ \frac{1}{N} \sum_{i=1}^{N} \overline{d}_i \le D,\ 0 \le y_1 - \overline{d}_1 \le \cdots \le y_N - \overline{d}_N \le T \right\}.$$

Equation (34) follows because $f_{\overline{D}}(\overline{d}_1, \ldots, \overline{d}_N) = \frac{1}{T^N}$. Let $\tilde{D} = (\tilde{D}_1, \ldots, \tilde{D}_N)$ represent a vector of independent, exponential random variables with common mean $\eta > 0$. The joint density of $\tilde{D}$ restricted to $\{\tilde{d} \in \Re_+^N | \frac{1}{N} \sum_{i=1}^N \tilde{d}_i \le D\}$ satisfies

$$f_{\tilde{D}}(\tilde{d}_1, \ldots, \tilde{d}_N) = \frac{1}{\eta^N} e^{-\frac{1}{\eta} \sum_{i=1}^N \tilde{d}_i} \ge \frac{1}{\eta^N} e^{-\frac{DN}{\eta}}.$$

Therefore, $\eta^N e^{\frac{DN}{\eta}} f_{\tilde{D}}(\tilde{d}_1, \ldots, \tilde{d}_N) \ge 1$ on this set, so (set $y_0 = 0$)

$$\begin{aligned} |A_y| &\le \int \cdots \int_{[0,T]^N} 1_{\{R\}} \eta^N e^{\frac{DN}{\eta}} \\ &\quad \cdot f_{\tilde{D}}\left(\overline{d}_1, \ldots, \overline{d}_N\right)\, d\overline{d}_1 \cdots d\overline{d}_N \\ &\le \eta^N e^{\frac{DN}{\eta}} P\begin{pmatrix} 0 \le y_1 - \tilde{D}_1 \le y_2 - \tilde{D}_2 \\ \le \cdots \le y_N - \tilde{D}_N \le T \end{pmatrix} \\ &\le \eta^N e^{\frac{DN}{\eta}} P\left(0 \le y_1 - \tilde{D}_1\right) \\ &\quad \cdot P\left(y_2 - \tilde{D}_2 \ge y_1 - \tilde{D}_1\right) \\ &\quad \cdots P\left(y_N - \tilde{D}_N \ge y_{N-1} - \tilde{D}_{N-1}\right) \qquad (35) \\ &= \eta^N e^{\frac{DN}{\eta}} \left(1 - e^{-y_1/\eta}\right) \prod_{i=1}^{N-1} \left(1 - \frac{1}{2} e^{-(y_{i+1} - y_i)/\eta}\right) \\ &\le \eta^N e^{\frac{DN}{\eta}}\, e^{\left[N\left(\frac{1}{N}\sum_{i=0}^{N-1} \log\left(1 - \frac{1}{2} e^{-(y_{i+1}-y_i)/\eta}\right)\right)\right]} \\ &\le \eta^N e^{\frac{DN}{\eta}}\, e^{\left[N \log\left(1 - \frac{1}{2} e^{-\frac{1}{N}\sum_{i=0}^{N-1} (y_{i+1}-y_i)/\eta}\right)\right]} \qquad (36) \\ &\le \eta^N e^{\frac{DN}{\eta}}\, e^{\left[N\left(\log\left(1 - \frac{1}{2} e^{-\frac{T}{N\eta}}\right)\right)\right]} \qquad (37) \end{aligned}$$

where (35) follows because the event $y_i - \tilde{D}_i \ge y_{i-1} - \tilde{D}_{i-1}$ can be rewritten as $E_1 = (\tilde{D}_i \le y_i - y_{i-1} + \tilde{D}_{i-1})$ and the event

$$\begin{aligned} E_2 = \{&0 \le y_1 - \tilde{D}_1, \\ &y_1 - \tilde{D}_1 \le y_2 - \tilde{D}_2, \ldots, y_{i-2} - \tilde{D}_{i-2} \le y_{i-1} - \tilde{D}_{i-1}\} \end{aligned}$$

simply implies a random upper bound on $\tilde{D}_{i-1}$ that is independent of $\tilde{D}_i$ and $\tilde{D}_{i-1}$ so that $P(E_1|E_2) \le P(E_1)$. Equation (36) follows by Jensen's Inequality because $\log(1 - \frac{1}{2} e^{-x/d})$ is concave in x and (37) follows because $\sum_{i=0}^{N-1} (y_{i+1} - y_i) \le T$. Since this bound does not depend on y we have that

$$\begin{aligned} H(X|Y) &\le \max_{\{y \in S\}} \log |A_y| \\ &\le \log\left(\eta^N e^{\frac{DN}{\eta}} \left(1 - \frac{1}{2} e^{-\frac{T}{N\eta}}\right)^N\right). \end{aligned}$$

Therefore, we have that

$$\begin{aligned} I(X;Y) &\ge \log \frac{T^N}{N!} - \log\left(\eta^N e^{\frac{DN}{\eta}} \left(1 - \frac{1}{2} e^{-\frac{T}{N\eta}}\right)^N\right) \\ &= \log \frac{T^N}{N! \left(\eta^N e^{\frac{DN}{\eta}} \left(1 - \frac{1}{2} e^{-\frac{T}{N\eta}}\right)^N\right)}. \end{aligned}$$

Using a Stirling bound ($n! < \sqrt{2\pi}\, n^{n+1/2} e^{-n+1/(12n)}$), we have that

$$I(X;Y) \geq N \log \left(\frac{T}{N\eta \exp \frac{D}{\eta} - 1 \left(1 - \frac{1}{2} e^{-\frac{T}{ND}}\right)} \right) - \log \left(\sqrt{2\pi} N^{1/2} e^{1/(12N)} \right).$$

Substituting $N = \lfloor \lambda T/k \rfloor$ and taking appropriate limits yields (28). □

B. Proof of Theorem VI.2

The proof of Theorem VI.2 closely follows the proof of Theorem VI.1. Each batch is of size k and there are $N = \lfloor \lambda T/k \rfloor$ batches in the input X. The T slots are divided into T/k super-slots, where either no packets or a complete batch of k packets is transmitted in each super-slot. Only the transmission times of the batches are informative, and we write X_i, $1 \leq i \leq N$ for the super-slot in which the ith batch is transmitted.

Let Y represent an output process for input X and a jammer in $\mathcal{W}_T$. Since packets are sent in complete batches, it is to the jammer's advantage to keep packets in the batch together while respecting super-slot boundaries. Without loss of generality, we assume that the output of the jammer is in that form since a trigger packetizer which groups packets into batches and a slotter which aligns batches with super-slot boundaries could be used at the jammer output. Similarly, since the decoder knows that all N batches are input to the jammer by time T, it is to the jammer's advantage to output all N batches by time T. Let Y_i represent the departure super-slot of the ith output batch. Thus, both X and Y are distributed in the set

$$S = \left\{ x \in Z^n \colon 1 \leq x_1 < x_2 < \cdots < x_N \leq \frac{T}{k} \right\}.$$

We will use the fact that $I(X;Y) = H(X) - H(X|Y)$. Since X is uniformly distributed over S, given any $\epsilon > 0$, if T is sufficiently large then

$$H(X) = \log \left(\binom{T/k}{N} \right) \geq \frac{T}{k} \left(h(\lambda) - \epsilon \right). \tag{38}$$

Given $y \in S$, let $A_y = \{x \in S \colon W_x = y\}$. Then $H(X|Y) \leq \max_y \log |A_y|$. For $y \in S$ and

$$R = \left\{ \frac{1}{N} \sum_{i=1}^{N} d_i \leq \frac{D}{k},\ 0 \leq y_1 - d_1 < \cdots < y_N - d_N \leq \frac{T}{k} \right\}$$

$$|A_y| = \sum_{d \in \{0, \ldots, \frac{T}{k} - 1\}^N} 1_{\{R\}}. \tag{39}$$

Given $0 < \eta < 1$, let $\tilde{D} = (\tilde{D}_1, \ldots, \tilde{D}_N)$, where the $\tilde{D}_i$'s are mutually independent and $P[\tilde{D}_i = j] = \eta(1-\eta)^j$ for $j \geq 0$. On the set $\{\frac{1}{N} \sum_{i=1}^{N} d_i \leq \frac{D}{k}\}$, the probability mass function $f_{\tilde{D}}$ satisfies

$$\eta^{-N} (1-\eta)^{-\frac{ND}{k}} f_{\tilde{D}}(d) = (1-\eta)^{d_1 + \cdots + d_N - \frac{ND}{k}} \geq 1. \tag{40}$$

Multiplying each term in (39) by the left-hand side of (40) yields

$$|A_y| \leq \eta^{-N} (1-\eta)^{-\frac{DN}{k}} \cdot P \left[1 \leq y_1 - \tilde{D}_1 < \cdots < y_N - \tilde{D}_N \leq \frac{T}{k} \right].$$

For $1 \leq k \leq N-1$

$$P \left[y_k - \tilde{D}_k < y_{k+1} - \tilde{D}_{k+1} \right] = \phi(y_{k+1} - y_k, \eta),$$

where $\phi(u, \eta) = 1 - \frac{(1-\eta)^u}{2-\eta}$. Also, $P[y_1 - \tilde{D}_1 \leq 1] \geq \phi(y_1, \eta)$. The function $\log \phi(u, \eta)$ is concave in u for $u > 0$, so arguing as in the proof of Theorem VI.1 yields

$$|A_y| \leq \left(\eta^{-1} (1-\eta)^{-\frac{D}{k}} \phi \left(\frac{T}{kN}, \eta \right) \right)^N. \tag{41}$$

Combining (38) and (41) yields that for T sufficiently large

$$\frac{I(X;Y)}{T} \geq \frac{h(\lambda)}{k} - \frac{\lambda}{k} \log \left(\eta^{-1} (1-\eta)^{-\frac{D}{k}} \phi \left(\frac{1}{\lambda}, \eta \right) \right) - \epsilon.$$

Taking appropriate limits yields (32), which can be optimized with respect to η and k. □

C. Jammers Satisfying Relaxed ADC Constraints

The particularly strict constraints considered for ADC jamming channels were chosen for technical reasons (for example, so that $C \leq \overline{V}$). Looser constraints for ADC jamming channels are that for any $T > 0$, inputs are in the class $\mathcal{X}_T = \{X \colon E[X(T)] = \lambda T\}$, where $X(T)$ is the number of packets up to time T and jammers are in the class

$$\mathcal{W}_T = \left\{ W \colon E \left[\frac{1}{X(T)} \sum_{i=1}^{X(T)} D_i \right] \leq D \right\}$$

where D_i is the delay added to the ith packet. These relaxed constraints allow us to consider a periodic dump jammer with period $2D$ as an ADC jamming channel with delay parameter D rather than being restricted to spacing D. In this appendix, we discuss jammers and inputs for these ADC jamming channels under these relaxed constraints.

As can be seen in Theorem IV.1, the maximum information rate for the periodic dump jammer with period $2D$ is given by $\frac{H(\mathrm{Geo}_0(2\lambda D))}{2D}$. The $\mathrm{Geo}_0(\lambda 2D)$-batch-with-spacing-$2D$ and the *compound-Poisson input* maximize the information through a periodic dump jammer with period $2D$. The compound-Poisson input transmits batches of packets at times chosen according to a Poisson process with rate λ, where the batch sizes are distributed according to

$$p_b(0) = 1 - \frac{\log(\lambda + 1)}{\lambda}$$

and

$$p_b(k) = \frac{1}{\lambda k} \left(\frac{\lambda}{\lambda + 1} \right)^k, \qquad k \in \{1, 2, \ldots\}$$

with a mean batch size of 1.

Another jamming channel is the *Poisson random ruler*, which is based on a collection of dump times chosen according to a Poisson process where all packets in the buffer at a dump time are released. In this setting, the arrival rate for the Poisson process is $\frac{1}{D}$ so that the average delay for packets passing through the jammer is D. An upper bound on the information rate for this jammer is $\frac{H(\mathrm{Geo}_0(\lambda D))}{D}$, which is also the maximum information rate for a periodic dump jammer with period D.

A memoryless server queue jammer satisfies the delay constraints assuming a Poisson input process. Anantharam and Verdú [3] discuss a memoryless server queue as a delay jammer. The drawback of this jammer is that there is no guarantee on

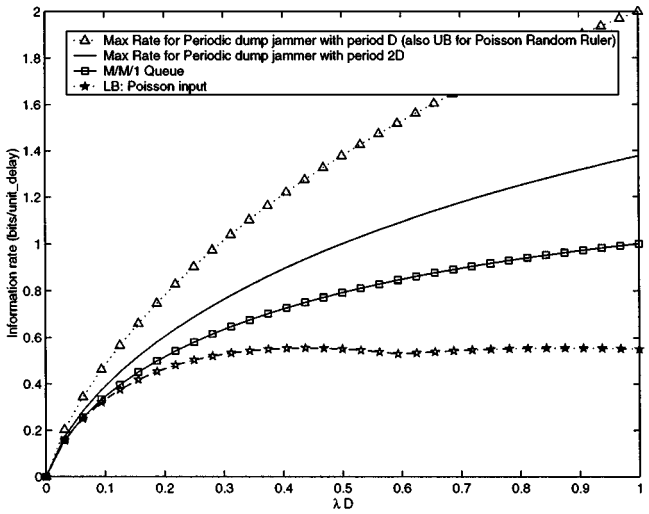

Fig. 22. Bounds on normalized information rate (bits/delay) for relaxed ADC jamming channels.

the maximum or average delay for most inputs. However, for the Poisson input, we can choose the service rate of the queue based on the arrival rate and the desired average delay. The information rate for the $M/M/1$ queue is given by $\lambda \log \frac{1+\lambda D}{\lambda D}$.

Under the relaxed ADC constraints, we can consider a Poisson input rather than the Poisson-like input. The lower bound for the Poisson input is the same in this case.

Bounds on $\min$-$\max$ and $\max$-$\min$ information rate for these jammers and inputs are plotted in Fig. 22. In particular, note the gap between the lower bound on the minimum information rate for the Poisson input, and the maximum rate for the periodic dump jammer with period $2D$. Assuming no duality gap, the value of the game is between these two bounds. Also note the difference between the periodic dump jammer with period $2D$ which can only be used for the relaxed ADC constraints, and the periodic dump jammer with period D which works for the regular ADC constraints.

REFERENCES

[1] R. Ahlswede and J. Wolfowitz, "Correlated decoding for channels with arbitrarily varying channel probability functions," *Inform. Contr.*, vol. 14, no. 5, pp. 457–473, May 1969.

[2] R. Ahlswede, "Elimination of correlation in random codes for arbitrarily varying channels," *Z. fur Wahrscheinlichkeitstheorie und verwandte Gebiete*, vol. 44, pp. 159–175, 1978.

[3] V. Anantharam and S. Verdú, "Bits through queues," *IEEE Trans. Inform. Theory*, vol. 42, pp. 4–18, Jan. 1996.

[4] A. S. Bedekar and M. Azizoglu, "The information-theoretic capacity of discrete-time queues," *IEEE Trans. Inform. Theory*, vol. 44, pp. 446–461, Mar. 1998.

[5] T. Berger, "Generalized billiard ball channels," in *Proc. 1999 IEEE Information Theory and Networking Workshop*, Metsovo, Greece, June 1999.

[6] T. Berger and Z. Zhang, "On the capacity of the grabbag channel," draft manuscript, May 2002.

[7] N. M. Blachman, "Communication as a game," in *1957 WESCON Conf. Rec.*, 1957, pp. 61–66.

[8] D. Blackwell, L. Breiman, and A. J. Thomasian, "Proof of Shannon's transmission theorem for finite-state indecomposable channels," *Ann. Math. Statist.*, vol. 31, pp. 558–867, 1958.

[9] O. L. Costich and I. S. Moskowitz, "Analysis of a storage channel in the two phase commit protocol," in *Proc. Computer Security Foundations Workshop IV*, June 1991, pp. 201–208.

[10] T. Cover and J. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[11] R. Cruz, "A calculus for network delay, Part I: Network elements in isolation," *IEEE Trans. Inform. Theory*, vol. 37, pp. 114–131, Jan. 1991.

[12] I. Csiszár and J. Körner, *Information Theory*. Budapest, Hungary: Akadémiai Kiadó, 1981.

[13] I. Csiszár and P. Narayan, "Arbitrarily varying channels with constrained inputs and states," *IEEE Trans. Inform. Theory*, vol. 34, pp. 27–34, Jan. 1988.

[14] ——, "Capacity and decoding rules for classes of arbitrarily varying channels," *IEEE Trans. Inform. Theory*, vol. 35, pp. 752–769, July 1989.

[15] R. L. Dobrushyn, "Optimal information transmission over a channel with unknown parameters," *Radiotekh. Elektron.*, vol. 4, no. 12, pp. 1951–1956, 1959.

[16] R. G. Gallager, "Basic limits on protocol information in data communication networks," *IEEE Trans. Inform. Theory*, vol. IT-22, pp. 385–398, July 1976.

[17] J. W. Gray, III, "On introducing noise into the bus-contention channel," in *Proc. 1993 IEEE Computer Society Symp. Research in Security and Privacy*, May 1993, pp. 90–98.

[18] M. V. Hegde, W. E. Stark, and D. Teneketzis, "On the capacity of channels with unknown interference," *IEEE Trans. Inform. Theory*, vol. 35, pp. 770–783, July 1989.

[19] W. Hu, "Reducing timing channels with fuzzy time," in *Proc. 1991 IEEE Computer Society Symp. Security and Privacy*, 1991, pp. 8–20.

[20] M. H. Kang and I. S. Moskowitz, "A data pump for communication," Naval Research Lab, Washington, DC, Tech. Rep. NRL/MR/55–95-7771, Sept. 1995.

[21] M. H. Kang, I. S. Moskowitz, and D. C. Lee, "A network pump," *IEEE Trans. Software Eng.*, vol. 22, pp. 329–337, May 1996.

[22] S. Karlin, *Mathematical Methods and Theory in Games, Programming and Economics*. Reading, MA: Addison-Wesley, 1959.

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

[24] A. Lapidoth and I. E. Telatar, "The compound channel capacity of a class of finite-state channels," *IEEE Trans. Inform. Theory*, vol. 44, pp. 973–983, May 1998.

[25] R. J. McEliece, "Communication in the presence of jamming—An information theory approach," in *Secure Digital Communications*. ser. CISM Courses and Lectures, G. Longo, Ed. New York: Springer-Verlag, 1983.

[26] P. M. Melliar-Smith and L. E. Moser, "Protection against covert storage and timing channels," in *Proc. Computer Security Foundations Workshop IV*, June 1991, pp. 209–214.

[27] J. K. Millen, "Finite-state noiseless covert channels," in *Proc. Computer Security Foundations Workshop II*, June 1989, pp. 81–86.

[28] I. S. Moskowitz, S. J. Greenwald, and M. H. Kang, "An analysis of the timed Z-channel," in *Proc. IEEE Symp. Security and Privacy*, May 1996, pp. 2–11.

[29] I. S. Moskowitz and M. H. Kang, "Discussion of a statistical channel," in *Proc. 1994 IEEE-IMS Workshop on Information Theory and Statistics*, October 1994, p. 95.

[30] ——, "Covert channels-here to stay?," in *Proc. COMPASS '94*, June 1994, pp. 235–243.

[31] I. S. Moskowitz and A. R. Miller, "Simple timing channels," in *Proc. 1994 IEEE Computer Society Symp. Research in Security and Privacy*, May 1994, pp. 56–64.

[32] National Computer Security Center, "A guide to understanding covert channel analysis of trusted systems," National Computer Security Center, Ft. George G. Meade, MD, Tech. Rep. NCSG-TG-030, Nov. 1993.

[33] S.-P. W. Shieh and V. D. Gligor, "Auditing the use of covert storage channels in secure systems," in *Proc. 1990 IEEE Computer Society Symp. Research in Security and Privacy*, May 1990, pp. 285–295.

[34] I. Stiglitz, "Coding for a class of unknown channels," *IEEE Trans. Inform. Theory*, vol. IT-12, pp. 189–195, Apr. 1966.

[35] R. Sundaresan and S. Verdú, "Robust decoding for timing channels," *IEEE Trans. Inform. Theory*, vol. 46, pp. 404–419, Mar. 2000.

[36] J. T. Trostle, "Multiple Trojan horse systems and covert channel analysis," in *Proc. Computer Security Foundations Workshop IV*, June 1991, pp. 22–33.

[37] C.-R. Tsai and V. D. Gligor, "A bandwidth computation model for covert storage channels and its applications," in *Proc. 1988 IEEE Symp. Security and Privacy*, Apr. 1988, pp. 108–121.

[38] B. R. Venkatraman and R. E. Wolfe, "Capacity estimation and auditability of network covert channels," in *Proc. 1995 IEEE Computer Society Symp. Security and Privacy*, May 1995, pp. 186–198.