Codes for Adversaries: Between Worst-Case and Average-Case Jamming

Other titles in Foundations and Trends® in Communications and Information Theory

Universal Features for High-Dimensional Learning and Inference
Shao-Lun Huang, Anuran Makur, Gregory W. Wornell and Lizhong Zheng
ISBN: 978-1-63828-176-4

Ultra-Reliable Low-Latency Communications: Foundations, Enablers, System Design, and Evolution Towards 6G
Nurul Huda Mahmood, Italo Atzeni, Eduard Axel Jorswieck and Onel Luis Alcaraz López
ISBN: 978-1-63828-180-1

Probabilistic Amplitude Shaping
Georg Böcherer
ISBN: 978-1-63828-178-8

Finite Blocklength Lossy Source Coding for Discrete Memoryless Sources
Lin Zhou and Mehul Motani
ISBN: 978-1-63828-182-5

Codes for Adversaries: Between Worst-Case and Average-Case Jamming

Bikash Kumar Dey
IIT Bombay
bikash@ee.iitb.ac.in

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

Michael Langberg
University at Buffalo
mikel@buffalo.edu

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

Yihan Zhang
IST Austria
zephyr.z798@gmail.com

now
the essence of knowledge
Boston — Delft

Foundations and Trends® in Communications and Information Theory

Published, sold and distributed by:
now Publishers Inc.
PO Box 1024
Hanover, MA 02339
United States
Tel. +1-781-985-4510
www.nowpublishers.com
sales@nowpublishers.com

Outside North America:
now Publishers Inc.
PO Box 179
2600 AD Delft
The Netherlands
Tel. +31-6-51115274

The preferred citation for this publication is

B. K. Dey *et al.*. *Codes for Adversaries: Between Worst-Case and Average-Case Jamming*. Foundations and Trends® in Communications and Information Theory, vol. 21, no. 3-4, pp. 300–588, 2024.

ISBN: 978-1-63828-461-1

Foundations and Trends® in Communications and Information Theory

Volume 21, Issue 3-4, 2024

Editorial Board

Editorial Scope

Foundations and Trends® in Communications and Information Theory publishes survey and tutorial articles in the following topics:

- Coded modulation
- Coding theory and practice
- Communication complexity
- Communication system design
- Cryptology and data security
- Data compression
- Data networks
- Demodulation and Equalization
- Denoising
- Detection and estimation
- Information theory and statistics
- Information theory and computer science
- Joint source/channel coding
- Modulation and signal design
- Multiuser detection
- Multiuser information theory
- Optical communication channels
- Pattern recognition and learning
- Quantization
- Quantum information processing
- Rate-distortion theory
- Shannon theory
- Signal processing for communications
- Source coding
- Storage and recording codes
- Speech and Image Compression
- Wireless Communications

Information for Librarians

Foundations and Trends® in Communications and Information Theory, 2024, Volume 21, 4 issues. ISSN paper version 1567-2190. ISSN online version 1567-2328 . Also available as a combined paper and online subscription.

Contents

Codes for Adversaries: Between Worst-Case and Average-Case Jamming

Bikash Kumar Dey[1], Sidharth Jaggi[2], Michael Langberg[3], Anand D. Sarwate[4] and Yihan Zhang[5]

[1] *Indian Institute of Technology Bombay, India; bikash@ee.iitb.ac.in*
[2] *University of Bristol, UK; sid.jaggi@bristol.ac.uk*
[3] *University at Buffalo, USA; mikel@buffalo.edu*
[4] *Rutgers University, USA; anand.sarwate@rutgers.edu*
[5] *Institute of Science and Technology Austria, Austria; zephyr.z798@gmail.com*

ABSTRACT

Over the last 70 years, information theory and coding has enabled communication technologies that have had an astounding impact on our lives. This is possible due to the match between encoding/decoding strategies and corresponding channel models. Traditional studies of channels have taken one of two extremes: Shannon-theoretic models are inherently average-case in which channel noise is governed by a memoryless stochastic process, whereas coding-theoretic (referred to as "Hamming") models take a worst-case, adversarial, view of the noise. However, for several existing and emerging communication systems the Shannon/average-case view may be too optimistic, whereas the Hamming/worst-case view may be too pessimistic. This monograph takes up

Bikash Kumar Dey, Sidharth Jaggi, Michael Langberg, Anand D. Sarwate and Yihan Zhang (2024), "Codes for Adversaries: Between Worst-Case and Average-Case Jamming", Foundations and Trends® in Communications and Information Theory: Vol. 21, No. 3-4, pp 300–588. DOI: 10.1561/0100000112.

the challenge of studying adversarial channel models that lie between the Shannon and Hamming extremes.

Preface

The arbitrarily varying channel (AVC) has often been considered an esoteric subject in information theory: a Shannon-theoretic take on worst-case communication that sometimes coincides with the notorious zero-error capacity problem. We think of the AVC and its many variants as capturing a class of models which are "between Shannon and Hamming" or between average-case and worst-case. We came to work on this topic via different routes but are motivated by a common question: what causes the gap between average-case and worst-case performance? It turns out there are many subtleties involved in reinvestigating the very basics of our communication models. As we dug deeper, we found new questions, even for the simplest of models, which required new techniques to answer.

This monograph is the product of research conducted by the authors and their collaborators over the last two decades. When we started writing it became clear that the task was more complicated than we had first imagined. A comprehensive treatment of the prior work on AVCs is necessary to understand the more recent models which form the later part of the monograph. The challenges of remote collaboration and the COVID pandemic stretched the process longer than we would have liked but we hope that you find it worth the wait!

Our goal in this work to convince the reader that there are fascinating connections between coding problems for AVCs and a wide range of topics ranging from game theory to tensor factorization to Ramsey theory. Tools such as list decoding and encoder randomization turn out to be natural tools for achievability arguments in these settings. List decoding is also a key tool in converse bounds, along with generalizations of the classical Plotkin bound. Giving a fresh look at these old topics can reveal interesting questions for future work.

1

Introduction

Information Theory and Coding Theory have made great advances since their start in the mid-20th century while having fundamentally different emphases. In the Shannon information-theoretic view, for general channels with *memoryless random* noise, we have codes that can achieve reliable communication at rates approaching capacity. At the other extreme, in the coding-theoretic view (which we refer to as "Hamming"), while some results are known for specific channels with *worst-case adversarial/malicious* noise (that may depend on the transmission), general capacity results are still elusive. The fundamental difference lies in the dependence between the message, the code, and the channel's effect on the transmitted symbols. Broadly speaking, Shannon-like models address average-case channel behavior whereas Hamming-like models deal with worst-case channel behavior.

This monograph addresses what happens in between these models. Recent work has identified a rich class of channels which interpolate between average and worst-case channel behavior. Using the language of arbitrarily varying channels (AVCs), these models consider a channel with a state input controlled by a malicious jammer who wishes to prevent reliable communication. The jammer's power comes in making the state dependent on the transmitted symbols and code structure. Several models that lie between worst-case and average-case behavior can be captured formally by characterizing the *jammer's view*: different views correspond to different types of dependence that the channel interference can have on the channel input.

Our study is motivated by fundamental issues in using information-theoretic models for communications as well as emerging applications for communication systems. In systems such as critical infrastructure and cyberphysical systems, low-power wireless communication systems for body-area networks, and multi-hop packet networks, the Shannon/average-case view may not be appropriate due to high variability in the channel, whereas the Hamming/worst-case view may be too pessimistic. Because the gap in capacity between these two models can be significant, it is important to understand where this gap comes from. As we will describe in this monograph, understanding these models that lie "between Shannon and Hamming" uncover some different insights about communication channels in terms of successful strategies for encoding/decoding and the nature of the "most harmful" interference.

1.1 The Jammer's View

At one extreme, there are jammers that can view the entire transmitted codeword noncausally before choosing an interference sequence; this corresponds to worst-case (Hamming-like) models. The other extreme includes jammers whose interference is oblivious of the transmitted codeword; corresponding the average-case (Shannon-type) models. Between these two extremes, one can consider a plethora of restrictions on the jammer's view. For example, consider the impact of *causality* and *myopia*: in the former, the jammer may be able to see transmitted symbols with some delay, while in the latter the jammer may observe noisy versions of the transmitted symbols. *How do these, and other, restrictions impact the achievable communication rate and code design?*

This monograph addresses models that lie between worst-case and average-case jamming behavior through the lens of the jammer's view. A *clear* view of the jammer brings it closer to the Hamming worst-case model, while an *obstructed* view moves the jammer towards the

Shannon average-case setting. In particular, the presentation highlights key mathematical tools, code construction strategies, and novel converse strategies for establishing capacity bounds and strict separations between the Shannon and Hamming models. While the monograph focuses on the impact on capacity of limiting the jammer's view of the transmitted codeword, we note that our perspective does not capture all aspects that differentiate between Shannon and Hamming type models, e.g., the memoryless nature of Shannon interference.

A driving force for several of the results presented in the monograph stems from new work on the worst-case Hamming model, which not only sheds new light for worst-case settings, but also advances knowledge in the intermediate models discussed throughout. Beyond the theoretical novelties presented here, this study is motivated by several existing and emerging communication systems. Applications of these models include smart infrastructures, autonomous vehicles, and other scenarios in which a random noise model is inappropriate but for which truly worst-case interference is too pessimistic.

1.2 Channel Modeling Using Arbitrarily Varying Channels (AVCs)

In what follows we present the basic model of study throughout the monograph. A more formal treatment appears in Section 2. We use the notation $[N] = \{1, 2, \ldots, N\}$. The starting point for our investigation is a channel with time-varying state. There are three parties in the model: a transmitter (Alice), a receiver (Bob), and a state generator (the jammer James). Alice wishes to send a message reliably to Bob over a channel that is partially controlled by James. For each channel use, James may generate a different state input. This model, which generalizes the compound channel, is known as an *arbitrarily varying channel* (AVC): a family of channels $\{W(y|x,s) : s \in \mathcal{S}\}$ with input x, output y, and state s taking values in the sets $\mathcal{X}$, $\mathcal{Y}$ and $\mathcal{S}$, respectively. We consider coding over a fixed blocklength n. The probability of an output sequence $\underline{y} \in \mathcal{Y}^n$, given an input sequence $\underline{x} \in \mathcal{X}^n$ and state sequence $\underline{s} \in \mathcal{S}^n$, is $W(\underline{y}|\underline{x},\underline{s}) = \prod_{t=1}^{n} W(y_t|x_t,s_t)$. An $(n, 2^{Rn})$ code for this channel is a pair of maps $(\mathsf{Enc}, \mathsf{Dec})$ where $\mathsf{Enc} : [2^{Rn}] \to \mathcal{X}^n$ and $\mathsf{Dec} : \mathcal{Y}^n \to [2^{Rn}]$. The maximal probability of error is defined as $\max_m \mathbb{P}(\mathsf{Dec}(\underline{y}) \neq m | \underline{x} = \mathsf{Enc}(m))$ and the average probability of error as $\frac{1}{2^{Rn}} \sum_{i=1}^{2^{Rn}} \mathbb{P}(\mathsf{Dec}(\underline{y}) \neq m | \underline{x} = \mathsf{Enc}(m))$. While many of our code construction use list-decoding (see Section 5), the criteria for successful communication is the standard one of *unique decoding*.

The common erasure and error models for binary-input channels can be cast in this framework by treating the erasure or error pattern as the state. In an *erasure model* we have $\mathcal{X}, \mathcal{S} = \{0, 1\}$ and $\mathcal{Y} = \{0, 1, \perp\}$, where $\perp$ stands for an erasure. The channel $W(y|x, 0)$ is noiseless and $y = x$ with probability 1. The channel $W(y|x, 1)$ is an erasure where $y = \perp$ with probability 1 regardless of x. The second model is a *bit-flip model* where $\mathcal{X}, \mathcal{S}, \mathcal{Y} = \{0, 1\}$ and the channel $W(y|x, s)$ satisfies $y = x \oplus s$ with probability 1, where $\oplus$ is addition modulo 2. These examples exhibit binary channels which are deterministic, however our framework supports a variety of models including channels over large (or continuous) alphabets and channels in which the action of James is governed by a general distribution $W(y|x, s)$ over y.

The codes described above are *deterministic* codes: each message m corresponds to a single codeword $\mathsf{Enc}(m)$. We also consider encoding functions $\mathsf{REnc}(m)$ using *private randomization*. These encoders are randomized map, but in privately randomized codes, the encoder randomness is known only by Alice and not revealed to Bob or James. This is in contrast to (fully) *randomized* encoding/decoding functions $\mathsf{REnc}(m)$ and $\mathsf{RDec}(m)$ where a source $\mathbf{r}$ of common randomness is shared by Alice and Bob, but not known to James [4], [31], [50], [51], [53], [103]. Randomized coding allows Alice and Bob to select a codebook privately without James's knowledge.[1] For both privately randomized and fully randomized codes we average the error probabilities (maximum and average) over encoder/shared randomness. When compared to privately randomized encoding, fully randomized coding gives Alice and Bob significantly more power. For example, if the common randomness is unlimited, they may *mask* the codeword (with a random permutation and additive *one time pad*), thus hiding the codeword completely from

[1] In some works on AVCs, codes with only private randomization are called "stochastic codes" and fully randomized codes are called "random codes." We are using more distinct terminology to help the reader remember the difference.

James. This typically reduces the Hamming setting to the Shannon one, as in the work of Bennett *et al.* [25]. The original AVC paper by Blackwell *et al.* [31] modeled communication as a game and randomized codes corresponded to *mixed strategies*. In this monograph we mainly focus on privately randomized codes and discuss fully randomized codes in Section 6.

1.3 Comparing Average-Case and Worst-Case Channel Behavior

1.3.1 Average case (Shannon)

The classical Shannon model [74], [149] for a discrete memoryless channel (DMC) or additive white Gaussian noise channel (AWGN) is shown in Figure 1.1.

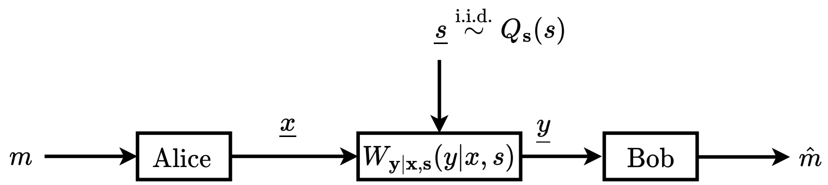

Figure 1.1: A memoryless channel with i.i.d. state. We can think of $\underline{s}$ as a state variable which contains the randomness in the channel. For additive channels such as the Binary Symmetric Channel or additive white Gaussian noise channel (AWGN) channel, the state can be taken as the noise in the channel.

The channel model assumes that the state sequence $\underline{s}$ is random and independent and identically distributed (i.i.d.) from some known distribution. A rate R is achievable if for any positive ε there exists a sufficiently large n and an $(n, 2^{Rn})$ code whose error is less than ε. The capacity is the supremum of achievable rates. For erasure (binary erasure channel $\mathsf{BEC}(p)$) and bit-flip (binary symmetric channel $\mathsf{BSC}(p)$) models the state $\underline{\mathbf{s}}$ is generated i.i.d. according to a Bernoulli distribution with parameter p. There are many strategies for achieving capacity in these channels, but the classical approach is *random coding* in which the codebook is constructed at random using a (single-letter) distribution over $\mathcal{X}$. For both the average and maximal error criteria, the capacities for the erasure and bit-flip models are $1 - p$ and $1 - H(p)$, respectively.

1.3.2 Worst-Case (Hamming)

The classical Hamming model [98] corresponds to the problem of error-control coding and is depicted in Figure 1.2. The state $\underline{s}$ controlled by James can be any sequence whose type belongs to a family Π_s of types over $\mathcal{S}$.

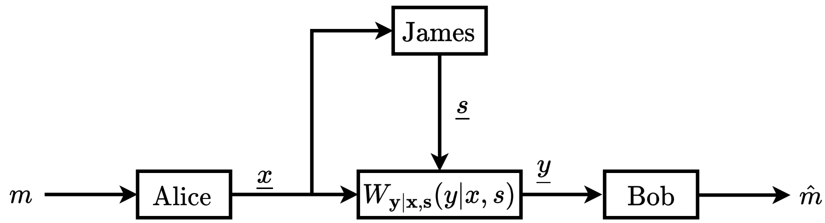

Figure 1.2: A channel with state controlled by an adversary. The state $\underline{s}$ is chosen to maximize the probability of decoding error. This is the model taken in classical coding theory.

A rate R is achievable if there is a sufficiently large n and an $(n, 2^{Rn})$ code with error equal to 0 (relaxing to a small positive average error does not change the achievable rate). For the erasure and bit-flip models, Π_s corresponds to sequences of $\{0, 1\}^n$ whose Hamming weight is at most pn.

The capacity in both the erasure and bit-flip models is upper bounded by the MRRW (or LP) bound [125] and lower bounded by the Gilbert-Varshamov (GV) bound [86], [165] (see Section 4 for more details in the erasure case).

1.3.3 Views between Shannon and Hamming

The Shannon and Hamming models represent two extremes: in the former, the state is i.i.d. and the goal is to achieve small error probability on *average* over interference. The Hamming model requires correct decoding for every $\underline{s}$ whose type lies in Π_s and represents a *worst-case* perspective. The traditional way to view this distinction is a difference in *error criterion*—the probability of error averaged over channel state versus the probability of error maximized over channel state. Here, we take a different perspective: we focus on *how the interference depends on* *the codeword*. In the Shannon model the state $\underline{s}$ is chosen independently of the codeword $\underline{x}$ whereas in the Hamming model the state can depend noncausally on the entire codeword $\underline{x}$.

To study models that lie between the Shannon and Hamming ones, we treat the state generator James as an *adversarial jammer* by explicitly describing how the state $\underline{s}$ can depend on the channel input $\underline{x}$ (and thus implicitly on the transmitted message m). We capture this dependence throughout the monograph by limiting James's view of the codeword (e.g., causal, myopic). Limiting James creates models between Shannon and Hamming: the stronger the limitations on James, the closer we are to the Shannon model. See Figure 1.3.

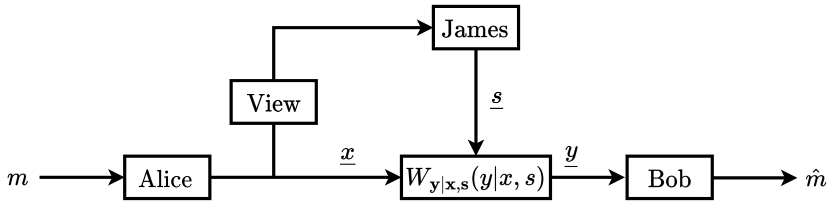

Figure 1.3: A channel with state that can depend on side information about the transmitted codeword $\underline{x}$. We model the state as controlled by a jammer James and call this side information the *jammer's view*. The view can be restricted in some way: for example, James may see a noisy version of $\underline{x}$ (myopic jammers) or be able to observe $\underline{x}$ sequentially (causal jammers).

1.3.4 Connection to arbitrarily varying channels

As discussed previously, the models studied herein are examples of the arbitrarily varying channel (AVC) model first proposed by Blackwell *et al.* [31]. The AVC model is broad in nature and, in its full generality, captures the setting in which the state vector $\underline{s}$ may depend on the transmitted codeword $\underline{x}$ and may be subject to lie in a given subset of $\mathcal{S}^n$. As such, the AVC model captures both the Hamming model and that of Shannon. Nevertheless, the majority of previous studies on AVCs address the Shannon model in which $\underline{s}$ does not depend on the transmitted codeword $\underline{x}$ [4], [50], [51], [53]. Early work focused on the difference between *randomized* and *deterministic* codes [4], [51]. In the randomized setting, Alice and Bob may mask the codeword and typically reduce the Hamming setting to that of Shannon [25], [31], [50], [81], [104], [117], [140], [141], [156], [160]. Ahlswede's classic derandomization technique [4] showed that the deterministic coding capacity of unconstrained AVCs is either 0 or equal to the randomized coding capacity. If James can "spoof" the codeword by selecting an input that makes the channel simulate a symmetric multiple access channel (MAC) with users Alice and James—the channel is *symmetrizable* and the capacity is 0 [51], [105]. We discuss the notion of symmetrizability in detail in later sections of the monograph. For more early results on AVCs, see Section 7 (also recommended is the excellent survey by Lapidoth *et al.* [119]).

1.4 Organization and Overview of Models Studied

This monograph is organized as follows. The first half of the monograph, including Section 2 through Section 5, sets the mathematical background and intuition towards the study of channels between the Shannon and Hamming models. Section 2 sets the notation and describes the models studied throughout the monograph. Section 3 and Section 4 present a number of motivating examples whose analyses are representative of those appearing later in the monograph. Section 5 presents a spectrum of results regarding list decoding in the context of AVCs. Although our ultimate goal in communication is that of unique decoding, as mentioned previously, list decoding, as a preliminary step in communication and as a measure of uncertainty of both the receiver Bob and jammer James, will play a major role in our analysis. The remaining sections of the monograph include a detailed analysis of the different channel models discussed in Section 2, including new results on general AVCs under the worst-case Hamming model; results which are used in the analysis of other models as well.

1.4.1 Section 3: Large Alphabets

The first example presented, studied in Section 3, addresses *causal adversarial models* in the setting in which the alphabet $\mathcal{X}$ of the codewords is *large*. Here, we investigate limitations on James through temporal constraints. Namely, James' action s_t at time t depends on his view of the codeword up to time $t-\Delta$. Equivalently, a codeword symbol transmitted at time t reaches James at time $t+\Delta$. Formally, we require for all t that $X^n \to X^{t-\Delta} \to S_t$ is a Markov chain. Delay $\Delta = n$ corresponds to the Shannon model, full lookahead $\Delta = -n$ corresponds to the Hamming model, and intermediate delay Δ bridges between these extremes. See Figure 1.4.

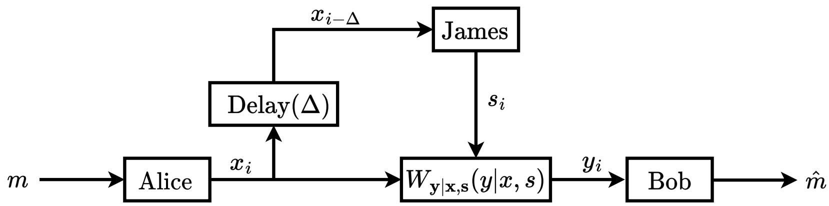

Figure 1.4: A channel with state that can depend on delayed observations of the transmitted codeword $\underline{x}$. The delay Δ controls how much of the codeword $\underline{x}$ James can see at each time i.

The study of causal jamming models in the large alphabet setting is an appropriate model for packet communication over multi-hop systems or ad-hoc networks in which a jammer can either eavesdrop or intercept transmissions over the channel. For example, in wireless packet communication, if James is eavesdropping, his action s_t at time t can depend only on past packets (i.e., $\Delta = 1$), whereas if he is acting as a relay he can tamper with the current channel input ($\Delta = 0$). The large alphabet setting studied in Section 3 allows us to reduce the causal adversarial model to the well understood model of erasures.

1.4.2 Section 4: Binary Erasures

Deviating from the large alphabet case, Section 4 studies both *causality* and *myopia* in the classical setting of binary channels. Here, the geometry of binary vs. large-alphabet vector spaces poses several challenges. To distill some of the main ideas, we focus on the simplest case of binary input channels - one with an adversary who can erase some of the transmitted bits. While causality deals with temporal constraints, myopia addresses interference in communication between Alice and the jammer James. Indeed, in several systems, such as all forms of wireless communication (over a private-designated or public-ISM band), there is little reason to assume that James has noiseless access to the transmitted codeword. This is true in any system operating in a noisy environment, from the emerging setting of IoT to that of Body-Area Networks. Thus, the study of myopic jammers arises naturally. More formally, myopic jammers can be modeled by an additional channel $W_{\mathbf{z}|\mathbf{x}}(z|x)$ between Alice and James. In this example section, we fix $W_{\mathbf{z}|\mathbf{x}}$ to be the binary erasure channel. See Figure 1.5.

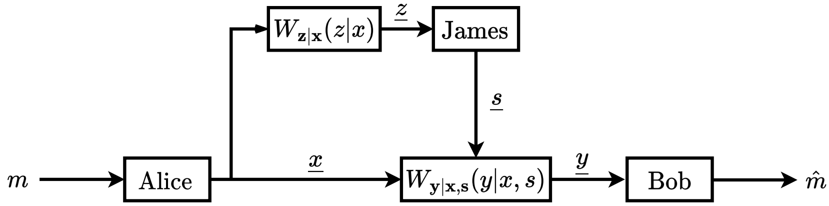

Figure 1.5: A channel with state that can depend on a noisy observation of the transmitted codeword $\underline{x}$. The channel $W(z|x)$ controls how good a view of the codeword $\underline{x}$ James has prior to his choosing $\underline{s}$.

A purely myopic jammer does not have additional causality constraints and communication can be considered to proceed in rounds. First Alice sends the encoding $\underline{x}$ over the channel, then James views the corrupted codeword $\underline{z}$ with probability $W^n(\underline{z}|\underline{x})$ (which equals $\prod_{t=1}^n W(z_t|x_t)$) and chooses a state sequence $\underline{s}$, finally Bob receives $\underline{y}$ with probability $\prod_{t=1}^n W(y_t|x_t,s_t)$ and decodes. The major challenges in the study of both the causal and myopic adversarial models are discussed. We review the major ideas and proof techniques to address these challenges as a preview to the upcoming sections.

1.4.3 Section 5: List Decoding

In list-decoding, one decodes a received word, not to a unique message, but rather to a list of potential messages. James's uncertainty about the transmitted codeword can be captured by the list of *potential codewords* transmitted by Alice that are consistent with James's view. Likewise, Bob's uncertainty is captured by the list consistent with his view. The evolution of James's and Bob's lists plays a central role in the design and analysis of coding schemes for the channel models studied. Quantifying the interplay between James and Bob in the communication process using the concept of list-decoding plays a major role in the monograph. Section 5 introduces a formal model for list decoding of AVCs in the average-case (Shannon) and worst-case (Hamming) settings. Quantitative bounds on the list size, decoding radii, and rate are discussed. The codes and results presented in Section 5 are used in the constructions for later sections.

1.4.4 Sections 6-11

The second half of the monograph, spanning Sections 6-11, includes a detailed discussion of the different channel models outlined in Section 2, starting from the more traditional Shannon and Hamming models of study to the newer models that lie between average- and worst-case analysis. For the traditional models, Section 6 starts with the study of AVCs in the setting of common randomness, Section 7 addresses oblivious AVCs for which the state s does not depend on the transmitted codeword $\underline{x}$, and Section 8 addresses the omniscient AVC setting in which James has full knowledge of the transmitted codeword. Myopic jammers, that view the transmitted codeword through a noisy channel, are studied in Section 9. Causal jammers, whose access of the transmitted codeword is limited by temporal constraints, are analyzed in Section 10. Finally, a collection of additional channel models between Shannon and Hamming are addressed in Section 11. A short description of the sections is given below.

- Section 6 reviews some of the "classical" results for AVCs with common randomness, starting with the first paper on AVCs [31] and then turning to methods for reducing the amount of common randomness for oblivious [4] and omniscient [117], [141], [156] adversaries - focusing on quantifying the amount of common randomness needed to achieve the randomized coding capacity. Here, the *oblivious* AVC model refers to the commonly studied setting in which James has no knowledge of the transmitted codeword $\underline{x}$.

- Section 7 gives a comprehensive study of the oblivious channel model in the setting in which common randomness is not permitted. From classical derandomization [4] to more general settings with constraints [51], the oblivious adversarial model without common randomness (the "standard" AVC model) has received a lot of attention in past decades. This section reviews the crucial notion of symmetrizability, which we later generalize for more complex adversarial settings.

- Section 8 revisits the Hamming setting for general AVCs and reviews the known bounds for positive zero-error capacity. A unified approach, via a geometric and effectively computable criterion [170], is presented for necessary and sufficient conditions for positive capacity. The sufficient condition presented leads to positive-rate code design via cloud codes, which are a strict generalization of Gilbert-Varshamov (GV) type codes. The necessary condition generalizes the Plotkin bound.

- Section 9 examines myopic jammers that access the codeword via a noisy channel. Central to the study of myopic jammers is the interplay between James' and Bob's view. *Does the jammer's side information reveal more information on the codeword transmitted than eventually available at the receiver Bob?* A jammer who can reveal more information than Bob is significantly more powerful than one who cannot. Governed by this dichotomy, the section reviews code design and converse proofs.

- Section 10 examines causal adversarial jammers that access and corrupt the codeword with temporal limitations. Tight achievability and converse proofs are given for a family of channel models.

- Finally, Section 11 touches briefly on several additional channel models and topics that fall within the general theme addressed by the monograph. These include, e.g., delayed jammers, quadratically constrained jammers, computational bounded jammers, jamming when the encoder possesses (noiseless) feedback, and more.

1.5 Major Analytical Tools and Techniques

In general, to leverage the limitations posed on James' view, it is crucial to design coding schemes that do not (implicitly) allow him to discover the transmitted codeword. More precisely, James should not be able to reliably choose a state vector $\underline{s}$ that results in a decoding error for Bob. For example, for deterministic additive channels $W(y|s,x)$ in which $y = x + s$, linear codes can at best achieve the (worst-case) Hamming capacity: the linear structure allows James to use the same state vector to cause an error for *every* codeword. Similarly, when using deterministic encoding functions, our model is of significance only under the average error criteria: under maximum error James need only cause an error on one message/codeword, which is exactly the Hamming model. With these challenges in mind, we here briefly outline the main analytical tools that allow to leverage James's limitations. Additional details are found in the sections that follow.

Privately randomized codes

In a privately randomized code, the randomness Alice uses in the encoding function $\mathsf{REnc}(m)$ is not known by Bob or James. Nevertheless, it can help: James has only limited knowledge of which codeword is transmitted, even if he knows the message m. Indeed, Ahlswede *et al.* [8] gave several equivalences between classes of AVC models and further showed that private randomization alone can have some benefit over deterministic encoding. Alice's ability to cause uncertainty at James through privately randomized coding is central to leverage the restrictions posed on James. However this comes at a price – Bob's uncertainty is simultaneously increased. Balancing the utility of privately randomized coding with this limiting factor plays a central role in the sections to come.

Chunkwise encoding

In our channel models, privately randomized codes need to be designed carefully to hide the transmitted codeword from James. For example, if the randomness of the encoder can be deduced by the jammer from a prefix or corrupted view of the codeword, then from that point on the code can be considered deterministic—the jammer holds full knowledge of the transmitted codeword. Therefore Alice must hide her random choices from the jammer, otherwise the setting is reduced to the worst-case setting of Hamming. The models studied throughout this monograph consider coding schemes in which encoder randomness is spread out *evenly* over the transmitted codeword, and cannot be deduced from limited views of the codeword. We call such schemes *chunkwise stochastic encoding schemes*. Formally, a chunkwise stochastic code of blocklength n consists of the concatenation of $\ell = \frac{n}{k}$ privately randomized codes of blocklength k, where each subcode typically uses independent randomness. Namely, $\mathsf{REnc}_n(m) = \mathsf{REnc}_k(m) \circ \mathsf{REnc}_k(m) \circ \cdots \circ \mathsf{REnc}_k(m)$, where each subcode uses independent encoder randomness. Here, the subscript n and k refer to the corresponding code's blocklegth. For example, chunkwise codes fit the temporal constraint of causal adversaries. If encoder randomness used in any codeword prefix is *independent* of that used in the remaining suffix, the jammer in his actions on the codeword prefix cannot *plan ahead* to fit the encoder randomness used in the design of the codeword suffix. This underlying structure is similar to block Markov encoding for relay channels [46], [75], [115], [164], [169], except that here the relay is the malicious adversary James, and can be used in studies of individual channels and streaming settings [82], [122], [152]. Chunkwise stochastic coding schemes leverage James's limitations and cause significant uncertainty in choosing the interference. However, the question now is how to deal with the increased uncertainty for Bob.

Converse proofs

The capacity gap between Shannon and Hamming models comes from understanding the jamming attacks that can be generated from limited adversaries. "Shannon-type" converse bounds, such as Fano's inequality, are too weak to model input-dependent interference, whereas combinatorial "Hamming-type" bounds, such as the Hamming, Singleton, or Plotkin bounds, strongly rely on James full knowledge of the transmitted codeword. This gives rise to the need of attacks that combine information-theoretic and combinatorial tools. For example, causal James may proceed in two phases by using "Shannon-like" input-independent interference in the first phase and then a "Hamming-like" input-dependent combinatorial attack in the second phase. The knowledge acquired by James in the first phase allows him to use input-dependent combinatorial bounds in the second. Examples of such attacks, termed "babble-and-push" attacks, appear throughout the monograph. Special emphasis is given on the concept of *symmetrizability*, which asks when James can make the channel at hand simulate a symmetric multiple access channel (MAC) with users Alice and James [51], [105], and as such cause an ambiguity about which message was encoded by which user. While symmetrizability is well understood in the oblivious setting, less is known in the models in which James holds (limited) codeword information.

1.6 A Note on Our Perspective

We wish to emphasize that our rhetorical use of "Shannon" and "Hamming" is not meant to imply that the particular channel modeling questions we discuss are the sole object of study in Shannon theory and coding theory. This monograph does not comprehensively cover all modeling options that lie between an average and worst-case error models. What we focus on are bounds on the capacity: finding the fundamental limits of achievable rates given the information available to the adversary. We will use random constructions to show bounds on the capacity and leave aside issues of computationally efficient code designs. We therefore do not examine more practical designs using sophisticated combinatorial and algebraic techniques that have been developed in coding theory. Our focus on point-to-point communication also not address the rich body of work on codes for other applications in which there are additional constraints on the encoding schemes such as locality or low-cost repair. These new settings have led to several breakthroughs in recent years.

References

[1] D. Aggarwal, Y. Dodis, and S. Lovett, "Non-malleable codes from additive combinatorics," in *Proceedings of the Forty-Sixth Annual ACM Symposium on Theory of Computing*, pp. 774–783, New York, NY, USA, 2014. DOI: 10.1145/2591796.2591804.

[2] R. Ahlswede, "A note on the existence of the weak capacity for channels with arbitrarily varying channel probability functions and its relation to Shannon's zero-error capacity," *Annals of Mathematical Statistics*, vol. 41, no. 3, 1970, pp. 1027–1033. DOI: 10.2307/2239255.

[3] R. Ahlswede, "Channels with arbitrarily varying channel probability functions in the presence of noiseless feedback," *Zeitschrift für Wahrscheinlichkeit und verwandte Gebiete*, vol. 25, no. 3, 1973, pp. 239–252. DOI: 10.1007/BF00535895.

[4] R. Ahlswede, "Elimination of correlation in random codes for arbitrarily varying channels," *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, vol. 44, no. 2, 1978, pp. 159–175. DOI: 10.1007/BF00533053.

[5] R. Ahlswede, "Arbitrarily varying channels with states sequence known to the sender," *IEEE Transactions on Information Theory*, vol. 32, no. 5, Sep. 1986, pp. 621–629. DOI: 10.1109/TIT.1986.1057222.

[6] R. Ahlswede and N. Cai, "Two proofs of Pinsker's conjecture concerning arbitrarily varying channels," *IEEE Transactions on Information Theory*, vol. 37, no. 6, Nov. 1991, pp. 1647–1649. URL: 10.1109/18.104326.

[7] R. Ahlswede and N. Cai, "Arbitrarily varying multiple-access channels part I—Ericson's symmetrizability is adequate, Gubner's conjecture is true," *IEEE Transactions on Information Theory*, vol. 45, no. 2, Mar. 1999, pp. 742–749. DOI: 10.1109/18.749024.

[8] R. Ahlswede and J. Wolfowitz, "Correlated decoding for channels with arbitrarily varying channel probability functions," *Information and Control*, vol. 14, no. 5, 1969, pp. 457–473. DOI: 10.1016/S0019-9958(69)90157-0.

[9] R. Ahlswede and N. Cai, "The AVC with noiseless feedback and maximal error probability: A capacity formula with a trichotomy," in *Numbers, Information and Complexity*, I. Althöfer, N. Cai, G. Dueck, L. Khachatrian, M. S. Pinsker, A. Sárközy, I. Wegener, and Z. Zhang, Eds., Boston, MA, USA: Springer, 2000, pp. 151–176. DOI: 10.1007/978-1-4757-6048-4_15.

[10] R. Ahlswede, C. Deppe, and V. Lebedev, "Non-binary error correcting codes with noiseless feedback, localized errors, or both," in *2006 IEEE International Symposium on Information Theory (ISIT)*, pp. 2486–2487, Seattle, WA, USA, Jul. 2006. DOI: 10.1109/ISIT.2006.262057.

[11] S. Arimoto, "An algorithm for computing the capacity of arbitrary discrete memoryless channels," *IEEE Transactions on Information Theory*, vol. 18, no. 1, Jan. 1972, pp. 14–20. DOI: 10.1109/TIT.1972.1054753.

[12] C. Baker and I. Chao, "Information capacity of channels with partially unknown noise. I. finite-dimensional channels," *SIAM Journal of Applied Mathematics*, vol. 56, no. 3, Jun. 1996, pp. 946–963. DOI: 10.1137/S0036139993249858.

[13] A. Barg, S. Guritman, and J. Simonis, "Strengthening the Gilbert-Varshamov bound," *Linear Algebra and its Applications*, vol. 307, no. 1-3, 2000, pp. 119–129. DOI: 10.1016/S0024-3795(99)00271-2.

[14] T. Başar, "The Gaussian test channel with an intelligent jammer," *IEEE Transactions on Information Theory*, vol. 29, no. 1, Jan. 1983, pp. 152–157. DOI: 10.1109/TIT.1983.1056602.

[15] T. Başar and Y. Wu, "A complete characterization of minimax and maximin encoder-decoder policies for communication channels with incomplete statistical description," *IEEE Transactions on Information Theory*, vol. 31, no. 4, Jul. 1985, pp. 482–489. DOI: 10.1109/TIT.1985.1057076.

[16] T. Ü. Başar and T. Başar, "Minimax causal transmission of Gaussian stochastic processes over channels subject to correlated jamming," in *Advances in Communications and Signal Processing*, ser. Lecture Notes in Information Sciences and Systems, W. Porter and S. Kak, Eds., vol. 129, Springer-Verlag, 1989, pp. 39–49. DOI: 10.1007/BFb0042717.

[17] T. Ü. Başar and T. Başar, "Optimum linear causal coding schemes for gaussian stochastic processes in the presence of correlated jamming," *IEEE Transactions on Information Theory*, vol. 35, no. 1, Jan. 1989, pp. 199–202. DOI: 10.1109/18.42196.

[18] T. Başar and Y. W. Wu, "Solutions to a class of minimax decision problems arising in communication systems," *Journal of Optimization Theory and Applications*, vol. 51, 1986, pp. 375–404. DOI: 10.1007/BF00940281.

[19] T. Ü. Başar and T. Başar, "Optimum coding and decoding schemes for the transmission of a stochastic process over a continuous-time stochastic channel with partially unknown statisticst," *Stochastics*, vol. 8, no. 3, 1982, pp. 213–237. DOI: 10.1080/17442508208833239.

[20] L. A. Bassalygo, "New upper bounds for error correcting codes," *Problemy Peredachi Informatsii*, vol. 1, no. 4, 1965, pp. 41–44. URL: https://www.mathnet.ru/eng/ppi762.

[21] R. Bassily and A. Smith, "Causal erasure channels," in *Proceedings of the Twenty-Fifth annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, Society for Industrial and Applied Mathematics, pp. 1844–1857, 2014. DOI: 10.1137/1.9781611973402.133.

[22] A. Bayesteh, M. Ansari, and A. Khandani, "Effect of jamming on the capacity of MIMO channels," in *42nd Annual Allerton Conference on Communication, Control, and Computing 2006*, pp. 401–410, Monticello, IL, USA: Curran Associates, Inc., 2004. URL: https://www.proceedings.com/00080.html.

[23] C. Bennett, P. Shor, J. Smolin, and A. Thapliyal, "Entanglement-assisted capacity of a quantum channel and the reverse Shannon theorem," *IEEE Transactions on Information Theory*, vol. 48, no. 10, Oct. 2002, pp. 2637–2655. DOI: 10.1109/TIT.2002.802612.

[24] C. H. Bennett, G. Brassard, C. Crepeau, and U. M. Maurer, "Generalized privacy amplification," *IEEE Transactions on Information Theory*, vol. 41, no. 6, Nov. 1995, pp. 1915–1923. DOI: 10.1109/18.476316.

[25] C. H. Bennett, G. Brassard, and J. Robert, "Privacy amplification by public discussion," *SIAM journal on Computing*, vol. 17, no. 2, 1988, pp. 210–229. DOI: 10.1137/0217014.

[26] C. H. Bennett, G. Brassard, and J. Robert, "Privacy amplification by public discussion," *SIAM Journal on Computing*, vol. 17, no. 2, 1988, pp. 210–229. DOI: 10.1137/0217014.

[27] E. R. Berlekamp, "Block coding with noiseless feedback," Ph.D. dissertation, Massachusetts Institute of Technology, Sep. 1964. URL: http://hdl.handle.net/1721.1/14783.

[28] A. Berman and N. Shaked-Monderer, *Completely Positive Matrices*. World Scientific, 2003. DOI: 10.1142/5273.

[29] N. Blachman, "Communication as a game," in *1957 IRE Wescon Conference Record*, vol. II, pp. 61–66, 1957.

[30] D. Blackwell, L. Breiman, and A. J. Thomasian, "The Capacity of a Class of Channels," *Annals of Mathematical Statistics*, vol. 30, no. 4, 1959, pp. 1229–1241.

[31] D. Blackwell, L. Breiman, and A. J. Thomasian, "The capacity of a certain channel classes under random coding," *Annals of Mathematical Statistics*, vol. 31, 1960, pp. 558–567. DOI: 10.1214/aoms/1177705783.

[32] R. Blahut, "Computation of channel capacity and rate-distortion functions," *IEEE Transactions on Information Theory*, vol. 18, no. 4, Jul. 1972, pp. 460–473. DOI: 10.1109/TIT.1972.1054855.

[33] V. M. Blinovsky, "Bounds for codes in the case of list decoding of finite volume," *Problems of Information Transmission*, vol. 22, no. 1, 1986, pp. 11–25. URL: https://www.mathnet.ru/eng/ppi839.

[34] I. M. Bomze, W. Schachinger, and R. Ullrich, "From seven to eleven: Completely positive matrices with high cp-rank," *Linear Algebra and its Applications*, vol. 459, Oct. 2014, pp. 208–221. DOI: 10.1016/j.laa.2014.06.025.

[35] J. M. Borden, D. M. Mason, and R. J. McEliece, "Some information theoretic saddlepoints," *SIAM Journal on Control and Optimization*, vol. 23, no. 1, 1985, pp. 129–143. DOI: 10.1137/0323011.

[36] A. J. Budkuley, B. K. Dey, S. Jaggi, M. Langberg, A. D. Sarwate, and C. Wang, "Symmetrizability for myopic AVCs," in *2020 IEEE International Symposium on Information Theory (ISIT)*, pp. 2103–2107, Los Angeles, CA, USA, Jun. 2020. DOI: 10.1109/ISIT44484.2020.9174487.

[37] C. E. Shannon, R. G. Gallager, and E. R. Berlekamp, "Lower bounds to error probability for coding on discrete memoryless channels. part i," *Information and Control*, vol. 10, no. 1, Jan. 1967, pp. 65–103. DOI: 10.1016/S0019-9958(67)90052-6.

[38] C. E. Shannon, R. G. Gallager, and E. R. Berlekamp, "Lower bounds to error probability for coding on discrete memoryless channels. part ii," *Information and Control*, vol. 10, no. 5, May 1967, pp. 522–552. DOI: 10.1016/S0019-9958(67)91200-4.

[39] C. Cachin and U. M. Maurer, "Linking information reconciliation and privacy amplification," *Journal of Cryptology*, vol. 10, 1997, pp. 97–110. DOI: 10.1007/s001459900023.

[40] Z. Chen, S. Jaggi, and M. Langberg, "A characterization of the capacity of online (causal) binary channels," in *Proceedings of the Forty-Seventh Annual ACM Symposium on Theory of Computing*, pp. 287–296, Portland, Oregon, USA, Jun. 2015. DOI: 10.1145/2746539.2746591.

[41] Z. Chen, S. Jaggi, and M. Langberg, "The capacity of online (causal) q-ary error-erasure channels," in *2016 IEEE International Symposium on Information Theory (ISIT)*, pp. 915–919, Barcelona, Spain, Jul. 2016. DOI: 10.1109/ISIT.2016.7541432.

[42] Z. Chen, S. Jaggi, and M. Langberg, "The capacity of online (causal) q-ary error-erasure channels," *IEEE Transactions on Information Theory*, vol. 65, no. 6, Jun. 2019, pp. 3384–3411. DOI: 10.1109/TIT.2019.2898863.

[43] M. Cheraghchi and V. Guruswami, "Capacity of non-malleable codes," *IEEE Transactions on Information Theory*, vol. 62, no. 3, Mar. 2015, pp. 1097–1118. DOI: 10.1109/TIT.2015.2511784.

[44] M. Cheraghchi and V. Guruswami, "Non-malleable coding against bit-wise and split-state tampering," *Journal of Cryptology*, vol. 30, 2017, pp. 191–241. DOI: 10.1007/s00145-015-9219-z.

[45] D. Conlon, J. Fox, and B. Sudakov, "Recent developments in graph Ramsey theory," in *Surveys in Combinatorics 2015*, ser. London Mathematical Society Lecture Note Series, A. Czumaj, A. Georgakopoulos, D. Král, V. Lozin, and O. Pikhurko, Eds., vol. 424, Cambridge, UK: Cambridge University Press, 2015, ch. 2, pp. 49–118. DOI: 10.1017/CBO9781316106853.

[46] T. M. Cover and A. El Gamal, "Capacity theorems for the relay channel," *IEEE Transactions on Information Theory*, vol. 25, no. 5, Sep. 1979, pp. 572–584. DOI: 10.1109/TIT.1979.1056084.

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

[48] R. Cramer, Y. Dodis, S. Fehr, C. Padró, and D. Wichs, "Detection of algebraic manipulation with applications to robust secret sharing and fuzzy extractors," in *Advances in Cryptology – EUROCRYPT 2008*, ser. Lecture Notes in Computer Science, N. Smart, Ed., Berlin, Heidelberg: Springer, 2008, pp. 471–488. DOI: 10.1007/978-3-540-78967-3_27.

[49] I. Csiszár and J. Körner, "On the capacity of the arbitrarily varying channel for maximum probability of error," *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, vol. 57, no. 1, 1981, pp. 87–101. DOI: 10.1007/BF00533715.

[50] I. Csiszár and P. Narayan, "Arbitrarily varying channels with constrained inputs and states," *IEEE Transactions on Information Theory*, vol. 34, no. 1, Jan. 1988, pp. 27–34. DOI: 10.1109/18.2598.

[51] I. Csiszár and P. Narayan, "The capacity of the arbitrarily varying channel revisited : Positivity, constraints," *IEEE Transactions on Information Theory*, vol. 34, no. 2, Mar. 1988, pp. 181–193. DOI: 10.1109/18.2627.

[52] I. Csiszár and P. Narayan, "Capacity and decoding rules for classes of arbitrarily varying channels," *IEEE Transactions on Information Theory*, vol. 35, no. 4, Jul. 1989, pp. 752–769. DOI: 10.1109/18.32153.

[53] I. Csiszár and P. Narayan, "Capacity of the Gaussian arbitrarily varying channel," *IEEE Transactions on Information Theory*, vol. 37, no. 1, Jan. 1991, pp. 18–26. DOI: 10.1109/18.61125.

[54] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*, Second. Cambridge University Press, Cambridge, 2011. DOI: 10.1017/CBO9780511921889.

[55] I. Csiszár and G. Tusnády, "Information geometry and alternating minimization procedures," *Statistics and Decisions (Supplement Issue)*, vol. 1, 1984, pp. 205–237.

[56] P. W. Cuff, H. H. Permuter, and T. M. Cover, "Coordination capacity," *IEEE Transactions on Information Theory*, vol. 56, no. 9, Sep. 2010, pp. 4181–4206. DOI: 10.1109/TIT.2010.2054651.

[57] P. Cuff, "Distributed channel synthesis," *IEEE Transactions on Information Theory*, vol. 59, no. 11, Nov. 2013, pp. 7071–7096. DOI: 10.1109/TIT.2013.2279330.

[58] A. D. Wyner, "The wire-tap channel," *The Bell System Technical Journal*, vol. 54, no. 8, Oct. 1975, pp. 1355–1387. DOI: 10.1002/j.1538-7305.1975.tb02040.x.

[59] B. K. Dey, S. Jaggi, and M. Langberg, "Codes against online adversaries: Large alphabets," *IEEE Transactions on Information Theory*, vol. 59, no. 6, Jun. 2013, pp. 3304–3316. DOI: 10.1109/TIT.2013.2245717.

[60] B. K. Dey, S. Jaggi, and M. Langberg, "Sufficiently myopic adversaries are blind," in *2015 IEEE International Symposium on Information Theory (ISIT)*, pp. 1164–1168, Hong Kong, China, Jun. 2015. DOI: 10.1109/ISIT.2015.7282638.

[61] B. K. Dey, S. Jaggi, and M. Langberg, "Sufficiently myopic adversaries are blind," *IEEE Transactions on Information Theory*, vol. 65, no. 9, Sep. 2019, pp. 5718–5736. DOI: 10.1109/TIT.2019.2916590.

[62] B. K. Dey, S. Jaggi, M. Langberg, A. D. Sarwate, and C. Wang, "The interplay of causality and myopia in adversarial channel models," in *2019 IEEE International Symposium on Information Theory (ISIT)*, IEEE, pp. 1002–1006, Paris, France, Jul. 2019. DOI: 10.1109/ISIT.2019.8849568.

[63] B. K. Dey, S. Jaggi, M. Langberg, and A. D. Sarwate, "Improved upper bounds on the capacity of binary channels with causal adversaries," in *2012 IEEE International Symposium on Information Theory (ISIT)*, Cambridge, MA, USA, Jul. 2012. DOI: 10.1109/ISIT.2012.6284300.

[64] B. K. Dey, S. Jaggi, M. Langberg, and A. D. Sarwate, "Upper bounds on the capacity of binary channels with causal adversaries," *IEEE Transactions on Information Theory*, vol. 59, no. 6, Jun. 2013, pp. 3753–3763. DOI: 10.1109/TIT.2013.2245721.

[65] B. K. Dey, S. Jaggi, M. Langberg, and A. D. Sarwate, "A bit of delay is sufficient and stochastic encoding is necessary to overcome online adversarial erasures," in *2016 IEEE International Symposium on Information Theory (ISIT)*, pp. 880–884, Barcelona, Spain, Jul. 2016. DOI: 10.1109/ISIT.2016.7541425.

[66] B. K. Dey, M. Langberg, S. Jaggi, and A. D. Sarwate, "Coding against delayed adversaries," in *2010 IEEE International Symposium on Information Theory (ISIT)*, pp. 285–289, Austin, Texas, USA, Jun. 2010. DOI: 10.1109/ISIT.2010.5513325.

[67] N. Do, "Party problems and Ramsey theory," *Vinculum*, vol. 56, no. 2, 2019, pp. 18–19. URL: https://search.informit.org/doi/abs/10.3316/ielapa.330480083052022.

[68] R. L. Dobrushin, "Unified information-transmission schemes for discrete memoryless channels and messages with independent components," *Doklady Akademii Nauk SSSR*, vol. 148, no. 6, 1963, pp. 1245–1248. URL: https://www.mathnet.ru/eng/dan27618.

[69] R. L. Dobrushin and S. Z. Stambler, “Coding theorems for classes of arbitrarily varying discrete memoryless channels,” *Problems of Information Transmission*, vol. 11, no. 2, 1975, pp. 97–112. URL: https://www.mathnet.ru/eng/ppi/v11/i2/p3.

[70] D. Dolev, C. Dwork, and M. Naor, “Non-malleable cryptography,” in *Proceedings of the Twenty-Third Annual ACM Symposium on Theory of Computing*, pp. 542–552, New Orleans, Louisiana, USA, 1991. DOI: 10.1145/103418.103474.

[71] S. C. Draper and A. Sahai, “Noisy feedback improves communication reliability,” in *2006 IEEE International Symposium on Information Theory (ISIT)*, pp. 69–73, Seattle, WA, USA, Jul. 2006. DOI: 10.1109/ISIT.2006.261676.

[72] S. Dziembowski, T. Kazana, and M. Obremski, “Non-malleable codes from two-source extractors,” in *Advances in Cryptology – CRYPTO 2013*, ser. Lecture Notes in Computer Science 8043, R. Canetti and J. A. Garay, Eds., Berlin, Heidelberg: Springer, 2013, pp. 239–257. DOI: 10.1007/978-3-642-40084-1_14.

[73] S. Dziembowski, K. Pietrzak, and D. Wichs, “Non-malleable codes,” *Journal of the ACM*, vol. 65, no. 4, 2018, pp. 1–32. DOI: 10.1145/3178432.

[74] C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication*. Urbana, IL: University of Illinois Press, 1949.

[75] A. El Gamal, N. Hassanpour, and J. Mammen, “Relay networks with delays,” *IEEE Transactions on Information Theory*, vol. 53, no. 10, Oct. 2007, pp. 3413–3431. DOI: 10.1109/TIT.2007.904838.

[76] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge, UK: Cambridge University Press, 2011.

[77] M. Elia, “Some results on the existence of binary linear codes,” *IEEE Transactions on Information Theory*, vol. 29, no. 6, 1983, pp. 933–934. DOI: 10.1109/TIT.1983.1056743.

[78] P. Elias, “List decoding for noisy channels,” in *Wescon Convention Record, Part 2*, pp. 94–104, Institute of Radio Engineers (now IEEE), 1957. URL: http://18.7.29.232/handle/1721.1/4484.

[79] P. Elias, “Error-correcting codes for list decoding,” *IEEE Transactions on Information Theory*, vol. 37, no. 1, Jan. 1991, pp. 5–12. DOI: 10.1109/18.61123.

[80] P. Erdős and G. Szekeres, "A combinatorial problem in geometry," *Compositio Mathematica*, vol. 2, 1935, pp. 463–470. URL: http://www.numdam.org/item?id=CM_1935___2___463_0.

[81] T. Ericson, "Exponential error bounds for random codes in the arbitrarily varying channel," *IEEE Transactions on Information Theory*, vol. 31, no. 1, Jan. 1985, pp. 42–48. DOI: 10.1109/TIT.1985.1056995.

[82] K. Eswaran, A. D. Sarwate, A. Sahai, and M. Gastpar, "Zero-rate feedback can achieve the empirical capacity," *IEEE Transactions on Information Theory*, vol. 56, no. 1, Jan. 2010, pp. 25–39. DOI: 10.1109/TIT.2009.2034779.

[83] F. Fabris, "Sharpening the Gilbert-Varshamov bound in the finite case," *Journal of Discrete Mathematical Sciences and Cryptography*, vol. 4, no. 1, 2001, pp. 65–75. DOI: 10.1080/09720529.2001.10697920. eprint: https://doi.org/10.1080/09720529.2001.10697920.

[84] P. Gács and J. Körner, "Common information is far less than mutual information," *Problems of Control and Information Theory*, vol. 2, no. 2, 1973, pp. 119–162. URL: http://real-j.mtak.hu/id/eprint/7981.

[85] S. I. Gel'fand and M. S. Pinsker, "Coding for channel with random parameters," *Problems of Control and Information Theory*, vol. 9, no. 1, 1980, pp. 19–31.

[86] E. N. Gilbert, "A comparison of signalling alphabets," *Bell Systems Technical Journal*, vol. 31, 1952, pp. 504–522. DOI: 10.1002/j.1538-7305.1952.tb01393.x.

[87] J. A. Gubner, "On the deterministic-code capacity of the multiple-access arbitrarily varying channel," *IEEE Transactions on Information Theory*, vol. 36, no. 2, Mar. 1990, pp. 262–275. DOI: 10.1109/18.52472.

[88] J. A. Gubner, "State Constraints for the Multiple-Access Arbitrarily Varying Channel," *IEEE Transactions on Information Theory*, vol. 37, no. 1, 1991, pp. 27–31.

[89] J. A. Gubner and B. Hughes, "Nonconvexity of the Capacity Region of the Multiple-Access Arbitrarily Varying Channel Subject to Constraints," *IEEE Transactions on Information Theory*, vol. 41, no. 1, 1995, pp. 3–13.

[90] V. Guruswami, "List decoding from erasures: bounds and code constructions," *IEEE Transactions on Information Theory*, vol. 49, no. 11, 2003, pp. 2826–2833. DOI: 10.1109/TIT.2003.815776.

[91] V. Guruswami and A. Rudra, "Explicit capacity-achieving list-decodable codes," in *Proceedings of the Thirty-Eighth Annual ACM Symposium on Theory of Computing (STOC '06)*, pp. 1–10, Seattle, WA, USA, 2006. DOI: 10.1145/1132516.1132518.

[92] V. Guruswami and M. Sudan, "Improved Decoding of Reed-Solomon and Algebraic-Geometry Codes," *IEEE Transactions on Information Theory*, vol. 45, no. 6, Sep. 1999, pp. 1757–1767. DOI: 10.1109/18.782097.

[93] V. Guruswami, *List Decoding of Error-Correcting Codes: Winning Thesis of the 2002 ACM Doctoral Dissertation Competition*, vol. 3282, ser. Lecture Notes in Computer Science. Springer, 2004. DOI: 10.1007/b104335.

[94] V. Guruswami, X. He, and R. Li, "The zero-rate threshold for adversarial bit-deletions is less than 1/2," *IEEE Transactions on Information Theory*, vol. 69, no. 4, Apr. 2022, pp. 2218–2239. DOI: 10.1109/TIT.2022.3223023.

[95] V. Guruswami, A. Rudra, and M. Sudan, "Essential coding theory." URL: https://cse.buffalo.edu/faculty/atri/courses/coding-theory/book/.

[96] V. Guruswami and A. Smith, "Codes for computationally simple channels: Explicit constructions with optimal rate," in *2010 IEEE 51st Annual Symposium on Foundations of Computer Science (FOCS)*, pp. 723–732, Las Vegas, NV, USA, 2010. DOI: 10.1109/FOCS.2010.74.

[97] V. Guruswami and C. Wang, "Linear-Algebraic List Decoding for Variants of Reed-Solomon Codes," *IEEE Transactions on Information Theory*, vol. 59, no. 6, Jun. 2013, pp. 3257–3268. DOI: 10.1109/TIT.2013.2246813.

[98] R. W. Hamming, "Error detecting and error correcting codes," *Bell Labs Technical Journal*, vol. 29, no. 2, 1950, pp. 147–160. DOI: 10.1002/j.1538-7305.1950.tb00463.x.

[99] A. Hashim, "Improvement on Varshamov-Gilbert lower bound on minimum Hamming distance of linear codes," *Proceedings of the Institution of Electrical Engineers*, vol. 125, no. 2, 1978, pp. 104–106. DOI: 10.1049/piee.1978.0028.

[100] M. Horstein, "Sequential transmission using noiseless feedback," *IEEE Transactions on Information Theory*, vol. 9, no. 3, Jul. 1963, pp. 136–143. DOI: 10.1109/TIT.1963.1057832.

[101] F. Hosseinigoki and O. Kosut, "The Gaussian interference channel in the presence of a malicious jammer," in *2016 54th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pp. 679–686, Monticello, IL, USA, 2016. DOI: 10.1109/ALLERTON.2016.7852297.

[102] F. Hosseinigoki and O. Kosut, "Capacity region of the Gaussian arbitrarily-varying broadcast channel," in *2020 IEEE International Symposium on Information Theory (ISIT)*, pp. 1007–1011, Los Angeles, CA, USA, Jun. 2020. DOI: 10.1109/ISIT44484.2020.9174108.

[103] B. Hughes and P. Narayan, "Gaussian arbitrarily varying channels," *IEEE Transactions on Information Theory*, vol. 33, no. 2, Mar. 1987, pp. 267–284. DOI: 10.1109/TIT.1987.1057288.

[104] B. L. Hughes and T. G. Thomas, "On error exponents for arbitrarily varying channels," *IEEE Transactions on Information Theory*, vol. 42, no. 1, Jan. 1996, pp. 87–98. DOI: 10.1109/18.481780.

[105] B. Hughes, "The smallest list for the arbitrarily varying channel," *IEEE Transactions on Information Theory*, vol. 43, no. 3, Mar. 1997, pp. 803–815. DOI: 10.1109/18.568692.

[106] S. Jaggi and M. Langberg, "Two-way interference channels with jammers," in *2017 IEEE International Symposium on Information Theory (ISIT)*, pp. 491–495, Aachen, Germany, Jun. 2017. DOI: 10.1109/ISIT.2017.8006576.

[107] J. Jahn, "Coding of arbitrarily varying multiuser channels," *IEEE Transactions on Information Theory*, vol. 27, no. 2, Mar. 1981, pp. 212–226. DOI: 10.1109/TIT.1981.1056320.

[108] T. Jiang and A. Vardy, "Asymptotic improvement of the Gilbert-Varshamov bound on the size of binary codes," *IEEE Transactions on Information Theory*, vol. 50, no. 8, Aug. 2004, pp. 1655–1664. DOI: 10.1109/TIT.2004.831751.

[109] A. Kashyap, T. Başar, and R. Srikant, "Correlated jamming on MIMO Gaussian fading channels," *IEEE Transactions on Information Theory*, vol. 50, no. 9, Sep. 2004, pp. 2119–2123. DOI: 10.1109/TIT.2004.833358.

[110] J. Kim, H. Liu, and T. Tran, "Exponential decay of intersection volume with applications on list-decodability and Gilbert-Varshamov type bound," *IEEE Transactions on Information Theory*, vol. 69, no. 5, May 2023, pp. 2841–2854. DOI: 10.1109/TIT.2022.3232241.

[111] J. Komlós, "A strange pigeon-hole principle," *Order*, vol. 7, no. 2, 1990, pp. 107–113. DOI: 10.1007/BF00383760.

[112] S. Kopparty, R. Shaltiel, and J. Silbak, "Quasilinear time list-decodable codes for space bounded channels," in *2019 IEEE 60th Annual Symposium on Foundations of Computer Science (FOCS)*, vol. 2019, pp. 302–333, Baltimore, MD, USA, 2019. DOI: 10.1109/FOCS.2019.00028.

[113] J. Körner and A. Orlitsky, "Zero-error information theory," *IEEE Transactions on Information Theory*, vol. 44, no. 6, 1998, pp. 2207–2229. DOI: 10.1109/18.720537.

[114] M. Kovačević and V. Y. F. Tan, "Codes in the space of multisets—coding for permutation channels with impairments," *IEEE Transactions on Information Theory*, vol. 64, no. 7, Jul. 2018, pp. 5156–5169. DOI: 10.1109/TIT.2017.2789292.

[115] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Transactions on Information Theory*, vol. 51, no. 9, 2005, pp. 3037–3063. DOI: 10.1109/TIT.2005.853304.

[116] R. La and V. Anantharam, "A game-theoretic look at the Gaussian multiaccess channel," in *DIMACS Workshop on Network Information Theory*, ser. DIMACS Series in Discrete Mathematics and Theoretical Computer Science, P. Gupta, G. Kramer, and A. van Winjgaarden, Eds., vol. 66, Piscataway, NJ: AMS DIMACS, Mar. 2003, pp. 87–106.

[117] M. Langberg, "Private codes or succinct random codes that are (almost) perfect," in *Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2004)*, Rome, Italy, 2004. DOI: 10.1109/FOCS.2004.51.

[118] M. Langberg, M. Schwartz, and E. Yaakobi, "Coding for the ℓ_∞ -limited permutation channel," *IEEE Transactions on Information Theory*, vol. 63, no. 12, Dec. 2017, pp. 7676–7686. DOI: 10.1109/TIT.2017.2762676.

[119] A. Lapidoth and P. Narayan, "Reliable communication under channel uncertainty," *IEEE Transactions on Information Theory*, vol. 44, no. 10, 1998, pp. 2148–2177. DOI: 10.1109/18.720535.

[120] A. Lapidoth and P. Narayan, "Reliable communication under channel uncertainty," *IEEE Transactions on Information Theory*, vol. 44, no. 6, 1998, pp. 2148–2177.

[121] C. T. Li and A. El Gamal, "An efficient feedback coding scheme with low error probability for discrete memoryless channels," *IEEE Transactions on Information Theory*, vol. 61, no. 6, Jun. 2015, pp. 2953–2963. DOI: 10.1109/TIT.2015.2428234.

[122] Y. Lomnitz and M. Feder, "Communication over individual channels," *IEEE Transactions on Information Theory*, vol. 57, no. 11, 2011, pp. 7333–7358. DOI: 10.1109/TIT.2011.2169130.

[123] A. Makur, "Coding theorems for noisy permutation channels," *IEEE Transactions on Information Theory*, vol. 66, no. 11, Nov. 2020, pp. 6723–6748. DOI: 10.1109/TIT.2020.3009468.

[124] U. Maurer and S. Wolf, "Privacy amplification secure against active adversaries," in *Advances in Cryptology — CRYPTO '97*, ser. Lecture Notes in Computer Science, B. S. Kaliski, Ed., vol. 1294, Berlin, Heidelberg: Springer Berlin Heidelberg, 1997, pp. 307–321. DOI: 10.1007/BFb0052244.

[125] R. J. McEliece, E. R. Rodemich, H. Rumsey Jr, and L. R. Welch, "New upper bounds on the rate of a code via the Delsarte-MacWilliams inequalities," *IEEE Transactions on Information Theory*, vol. 23, no. 2, 1977, pp. 157–166. DOI: 10.1109/TIT.1977.1055688.

[126] M. Médard, "Capacity of correlated jamming channels," in *Proceedings of the Annual Allerton Conference on Communication Control and Computing*, vol. 35, pp. 1043–1052, Monticello, IL, USA, 1997.

[127] M. Mitzenmacher, "A survey of results for deletion channels and related synchronization channels," *Probability Surveys*, vol. 6, 2009, pp. 1–33. DOI: 10.1214/08-PS141.

[128] P. Narayan and H. Tyagi, "Multiterminal secrecy by public discussion," *Foundations and Trends in Communications and Information Theory*, vol. 13, no. 2-3, 2016, pp. 129–275. DOI: 10.1561/0100000072.

[129] S. Nishimura, "The strong converse theorem in the decoding scheme of list size L," *Kōdai Mathematical Seminar Reports*, vol. 21, no. 4, 1969, pp. 418–425. DOI: 10.2996/kmj/1138845989.

[130] O. Øystein, "Über höhere kongruenzen," *Norsk Matematisk Forenings Skrifter Serie I*, vol. 7, 1922, p. 15.

[131] U. Pereg and Y. Steinberg, "The arbitrarily varying channel under constraints with side information at the encoder," *IEEE Transactions on Information Theory*, vol. 65, no. 2, Feb. 2019, pp. 861–887. DOI: 10.1109/TIT.2018.2861776.

[132] U. Pereg and Y. Steinberg, "The arbitrarily varying relay channel," *Entropy*, vol. 21, no. 516, 2019. DOI: 10.3390/e21050516.

[133] U. Pereg and Y. Steinberg, "The arbitrarily varying broadcast channel with causal side information at the encoder," *IEEE Transactions on Information Theory*, vol. 66, no. 2, 2020, pp. 757–779. DOI: 10.1109/TIT.2019.2927696.

[134] M. Plotkin, "Binary codes with specified minimum distance," *IRE Transactions on Information Theory*, vol. 6, no. 4, 1960, pp. 445–450. DOI: 10.1109/TIT.1960.1057584. (accessed on 10/08/2015).

[135] I. Reed and G. Solomon, "Polynomial Codes Over Certain Finite Fields," *Journal of the Society for Industrial and Applied Mathematics*, vol. 8, no. 2, 1960, pp. 300–304. DOI: 10.1137/0108018.

[136] A. Rudra and S. Uurtamo, "Two theorems on list decoding," in *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, M. Serna, R. Shaltiel, K. Jansen, and J. Rolim, Eds., ser. Lecture Notes in Computer Science, vol. 6302, pp. 696–709, Berlin, Heidelberg: Springer Berlin Heidelberg, 2010. DOI: 10.1007/978-3-642-15369-3_52.

[137] E. Ruzomberka, C. Wang, and D. J. Love, "Channel capacity for adversaries with computationally bounded observations," in *2022 IEEE International Symposium on Information Theory (ISIT)*, vol. abs/2202.02905, pp. 2529–2534, Espoo, Finland, Jun. 2022. DOI: 10.1109/ISIT50566.2022.9834532. arXiv: 2202.02905.

[138] A. D. Sarwate, "Robust and adaptive communication under uncertain interference," Ph.D. dissertation, University of California, Berkeley, Jul. 2008. URL: https://www.eecs.berkeley.edu/Pubs/TechRpts/2008/EECS-2008-86.pdf.

[139] A. D. Sarwate, "Coding against myopic adversaries," in *Proceedings of the 2010 Information Theory Workshop*, pp. 1–5, Dublin, Ireland, Aug. 2010. DOI: 10.1109/CIG.2010.5592896.

[140] A. D. Sarwate and M. Gastpar, "Randomization bounds on Gaussian arbitrarily varying channels," in *2006 IEEE International Symposium on Information Theory (ISIT)*, pp. 2161–2165, Seattle, WA, USA, Jul. 2006. DOI: 10.1109/ISIT.2006.261933.

[141] A. D. Sarwate and M. Gastpar, "Rateless codes for AVC models," *IEEE Transactions on Information Theory*, vol. 56, no. 7, Jul. 2010, pp. 3105–3114. DOI: 10.1109/TIT.2010.2048497.

[142] S. Satpathy and P. Cuff, "Secure cascade channel synthesis," in *2013 IEEE International Symposium on Information Theory (ISIT)*, pp. 2955–2959, Istanbul, Turkey, Jul. 2013. DOI: 10.1109/ISIT.2013.6620767.

[143] J. P. M. Schalkwijk and T. Kailath, "A coding scheme for additive noise channels with feedback—part I: No bandwidth constraint," *IEEE Transactions on Information Theory*, vol. 12, no. 2, Apr. 1966, pp. 172–182. DOI: 10.1109/TIT.1966.1053879.

[144] J. T. Schwartz, "Fast probabilistic algorithms for verification of polynomial identities," *Journal of ACM*, vol. 27, no. 4, 1980, pp. 701–717. DOI: 10.1145/322217.322225.

[145] S. Shafiee and S. Ulukus, "Mutual information games in multiuser channels with correlated jamming," *IEEE Transactions on Information Theory*, vol. 55, no. 10, Oct. 2009, pp. 4598–4607. DOI: 10.1109/TIT.2009.2027577.

[146] R. Shaltiel and J. Silbak, "Explicit list-decodable codes with optimal rate for computationally bounded channels," in *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2016)*, K. Jansen, C. Mathieu, J. D. P. Rolim, and C. Umans, Eds., ser. Leibniz International Proceedings in Informatics (LIPIcs), vol. 60, 45:1–45:38, Dagstuhl, Germany: Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2016. DOI: 10.4230/LIPIcs.APPROX-RANDOM.2016.45.

[147] R. Shaltiel and J. Silbak, "Explicit uniquely decodable codes for space bounded channels that achieve list-decoding capacity.," in *STOC 2021: Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*, vol. 27, pp. 1516–1526, 2021. DOI: 10.1145/3406325.3451048.

[148] R. Shaltiel and J. Silbak, "Error correcting codes that achieve BSC capacity against channels that are poly-size circuits," in *022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)*, pp. 13–23, Denver, CO, USA, 2022. URL: 10.1109/FOCS54457.2022.00009.

[149] C. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, no. 3, Jul. 1948, pp. 379–423, 623–656. DOI: 10.1002/j.1538-7305.1948.tb01338.x.

[150] C. E. Shannon, "The zero error capacity of a noisy channel," *IRE Transactions on Information Theory*, vol. 2, no. 3, Sep. 1956, pp. 8–19. DOI: 10.1109/TIT.1956.1056798.

[151] C. E. Shannon, "Communication Theory of Secrecy Systems," *The Bell System Technical Journal*, vol. 28, no. 4, Oct. 1949, pp. 656–715. DOI: 10.1002/j.1538-7305.1949.tb00928.x.

[152] O. Shayevitz and M. Feder, "Achieving the empirical capacity using feedback: Memoryless additive models," *IEEE Transactions on Information Theory*, vol. 55, no. 3, 2009, pp. 1269–1295. DOI: 10.1109/TIT.2008.2011434.

[153] O. Shayevitz and M. Feder, "Optimal feedback communication via posterior matching," *IEEE Transactions on Information Theory*, vol. 57, no. 3, Mar. 2011, pp. 1186–1222. DOI: 10.1109/TIT.2011.2104992.

[154] R. C. Singleton, "Maximum distance q-nary codes," *IEEE Transactions on Information Theory*, vol. 10, no. 2, Apr. 1964, pp. 116–118. DOI: 10.1109/TIT.1964.1053661.

[155] M. Sion, "On general minimax theorems," *Pacific Journal of Mathematics*, vol. 8, no. 1, 1958, pp. 171–176. URL: http://projecteuclid.org/euclid.pjm/1103040253.

[156] A. Smith, "Scrambling adversarial errors using few random bits, optimal information reconciliation, and better private codes," in *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 2007)*, pp. 395–404, New Orleans, LA, USA, Jan. 2007. URL: https://eprint.iacr.org/2006/020.pdf.

[157] W. E. Stark and R. J. McEliece, "On the capacity of channels with block memory," *IEEE Transactions on Information Theory*, vol. 34, no. 2, 1988, pp. 322–324. DOI: 10.1109/18.2642.

[158] M. Sudan, "Decoding of Reed-Solomon codes beyond the error-correction bound," *Journal of Complexity*, vol. 13, no. 1, Mar. 1997, pp. 180–193. DOI: 10.1006/jcom.1997.0439.

[159] J. Tang and Y. Polyanskiy, "Capacity of noisy permutation channels," in *2022 IEEE International Symposium on Information Theory (ISIT)*, pp. 1987–1992, Espoo, Finland, Jun. 2022. DOI: 10.1109/ISIT50566.2022.9834509.

[160] T. G. Thomas and B. Hughes, "Exponential error bounds for random codes on Gaussian arbitrarily varying channels," *IEEE Transactions on Information Theory*, vol. 37, no. 3, May 1991, pp. 643–649. DOI: 10.1109/18.79922.

[161] P. Tian, S. Jaggi, M. Bakshi, and O. Kosut, "Arbitrarily varying networks: Capacity-achieving computationally efficient codes," in *2016 IEEE International Symposium on Information Theory (ISIT)*, pp. 2139–2143, Barcelona, Spain, Jul. 2016. DOI: 10.1109/ISIT.2016.7541677.

[162] L. Tolhuizen, "The generalized Gilbert-Varshamov bound is implied by Turán's theorem [code construction]," *IEEE Transactions on Information Theory*, vol. 43, no. 5, 1997, pp. 1605–1606. DOI: 10.1109/18.623158.

[163] M. A. Tsfasman, S. G. Vlăduţ, and T. Zink, "Modular curves, Shimura curves, and Goppa codes, better than Varshamov-Gilbert bound," *Mathematische Nachrichten*, vol. 109, no. 1, 1982, pp. 21–28. DOI: 10.1002/mana.19821090103.

[164] E. C. van der Meulen, "Three-terminal communication channels," *Advances in Applied Probability*, vol. 3, no. 1, 1971, pp. 120–154. DOI: 10.1017/S0001867800037605.

[165] R. R. Varshamov, "The evaluation of signals in codes with correction of errors," *Doklady Akademii Nauk SSSR*, vol. 117, no. 5, 1957, pp. 739–741. URL: https://www.mathnet.ru/eng/dan22571.

[166] B. N. Vellambi, J. Kliewer, and M. R. Bloch, "Strong coordination over multi-hop line networks using channel resolvability codebooks," *IEEE Transactions on Information Theory*, vol. 64, no. 2, Feb. 2018, pp. 1132–1162. DOI: 10.1109/TIT.2017.2768529.

[167] V. Vu and L. Wu, "Improving the Gilbert-Varshamov bound for q-ary codes," *IEEE Transactions on Information Theory*, vol. 51, no. 9, 2005, pp. 3200–3208. DOI: 10.1109/TIT.2005.853300.

[168] C. Wang, "On the capacity of the binary adversarial wiretap channel," in *Proceedings of the 54th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pp. 363–369, Monticello, IL, USA, 2016. DOI: 10.1109/ALLERTON.2016.7852254.

[169] L. Wang and M. Naghshvar, "On the capacity of the noncausal relay channel," *IEEE Transactions on Information Theory*, vol. 63, no. 6, Jun. 2017, pp. 3554–3564. DOI: 10.1109/TIT.2017.2697868.

[170] X. Wang, A. J. Budkuley, A. Bogdanov, and S. Jaggi, "When are large codes possible for AVCs?" In *2019 IEEE International Symposium on Information Theory (ISIT)*, pp. 632–636, Paris, France: IEEE, Jul. 2019. DOI: 10.1109/ISIT.2019.8849324.

[171] H. Witsenhausen, "On sequences of pairs of dependent random variables," *SIAM Journal on Applied Mathematics*, vol. 28, no. 1, 1975, pp. 100–113. DOI: 10.1137/0128010.

[172] J. M. Wozencraft, "List decoding," in *Quarterly Progress Report*, 48, J. B. Wiesner, G. G. Harvey, and H. J. Zimmerman, Eds., MIT Reseearch Laboratory of Electronics, Jan. 1958, ch. XII, pp. 90–95. URL: http://hdl.handle.net/1721.1/52133.

[173] A. Wyner, "The common information of two dependent random variables," *IEEE Transactions on Information Theory*, vol. 21, no. 2, Mar. 1975, pp. 163–179. DOI: 10.1109/TIT.1975.1055346.

[174] A. K. Yadav, M. Alimohammadi, Y. Zhang, A. J. Budkuley, and S. Jaggi, "New results on AVCs with omniscient and myopic adversaries," in *2022 IEEE International Symposium on Information Theory (ISIT)*, IEEE, pp. 2535–2540, Espoo, Finland, Jun. 2022. DOI: 10.1109/ISIT50566.2022.9834632.

[175] Q. Zhang, S. Kadhe, M. Bakshi, S. Jaggi, and A. Sprintson, "Reliable and secure communication over adversarially jammed multipath networks, Part I: One-shot setting," 2022.

[176] Y. Zhang, S. Jaggi, and A. J. Budkuley, "Tight list-sizes for oblivious AVCs under constraints," ArXiV, Tech. Rep. 2009.03788 [cs.IT], Sep. 2020. DOI: 10.48550/arXiv.2009.03788.

[177] Y. Zhang, S. Jaggi, M. Langberg, and A. D. Sarwate, "The capacity of causal adversarial channels," in *2022 IEEE International Symposium on Information Theory (ISIT)*, pp. 2523–2528, Espoo, Finland, Jun. 2022. DOI: 10.1109/ISIT50566.2022.9834709.

[178] Y. Zhang, S. Jaggi, M. Langberg, and A. D. Sarwate, "The capacity of causal adversarial channels," ArXiV, Tech. Rep. 2205.06708 [cs.IT], May 2022. DOI: 10.48550/arXiv.2205.06708.

[179] Y. Zhang, S. Vatedka, and S. Jaggi, "Quadratically constrained two-way adversarial channels," in *2020 IEEE International Symposium on Information Theory (ISIT)*, IEEE, pp. 1587–1592, Los Angeles, CA, USA, Jun. 2020. DOI: 10.1109/ISIT44484.2020.9174421.

[180] K. S. Zigangirov, "On the number of correctable errors for transmission over a binary symmetrical channel with feedback," *Problemy Peredachi Informatsii*, vol. 12, no. 2, 1976, pp. 3–19. URL: https://www.mathnet.ru/eng/ppi1683.

[181] R. Zippel, "Probabilistic algorithms for sparse polynomials," in *EUROSAM 1979: Symbolic and Algebraic Computation*, ser. Lecture Notes in Computer Science, E. W. Ng, Ed., vol. 72, Berlin, Heidelberg: Springer, 1979. DOI: 10.1007/3-540-09519-5_73.