IEEE Information Theory Society Newsletter

Vol. 69, No. 3, September 2019 EDITOR: Salim El Rouayheb ISSN 1059-2362

President's Column

Emina Soljanin

The International Symposium on Information Theory (ISIT), our Society's flagship conference, took place a few weeks ago in Paris. It was there, in *la ville-lumière*, that many of us were able to discern some light at the end of the tunnel we felt we entered more than a year ago. But some have told me that they did not think we got completely out of the (Vail) woods yet. I am now at home, looking at the skyline of Manhattan, *the city that never sleeps*, and reflecting back on the ISIT and the time from the beginning of the year.

The ISIT was a great success by all measures. (Alright, I know, many would use different words for the banquet, and I will come back to that.) We had a record number of attendees and a very exciting plenary and regular program. We had five special sessions on information theory and related fields, as a part of this year's new initiative I wrote about in the March issue. The membership events were, as always, well attended and well received. We had a superb Shannon lecture, delivered flawlessly by Erdal Arıkan followed by a very lively discussion.

The annual Society's awards, listed later in this issue, were presented at the ISIT award ceremony. Our small Society (by the IEEE membership measure), as usual, received a disproportionately large number of IEEE level recognitions. This year, we were honored to have the IEEE President José Moura present some of these awards to our members.

The 2020 Shannon Award, our Society's most prestigious recognition, went to Charles Bennett, a researcher at IBM, New York. As nicely stated by Andreas Winter, one of Bennett's collaborators, *Charles Bennett has been instrumental in the creation of modern quantum information from the 1980s. Even if he didn't create the new information science on his own, Charlie had his hand in every fundamental conceptual breakthrough, and they all show his signature thinking about information, which is informed by his interest in the physical representation of information and by his unique way of looking at information theory as a physical theory.*

You may think that quantum information theory is an exotic area, and at most an esoteric subfield of information theory. Nevertheless, it is classical information theory that is a special (non-contextual) case of quantum information theory. You may want to check with Google in which class of papers the term "information theory" is used the most often nowadays. You may be surprised. It has been alleged that the mathematician David Hilbert had said that *physics is too important to be left to the physicists* to which the physicist John Wheeler, several decades later in retaliation, responded that *Gödel is too important to be left to the mathematicians* [1]. Today, I dare say, they would both agree that information theory is too important to be left to the information theorists. And so would many contemporary physicists, mathematicians, and computer scientists.

Now back to the banquet. I was disappointed too. I had a little speech, which could not be heard because of the acoustics in the space. I had prepared jokes and hints to help the audience guess the 2020 Shannon award winner's identity, which I did not have a chance to say. But, what I was truly sorry about was the lost opportunity for our society members to *sit* together and enjoy a well deserved rest and food. That would have provided a further chance for our Society to heal.

Yet, I realize how much work it takes to organize an ISIT. So much so that I never ventured into serving as a general chair to one, in spite of having extensive experience in organizing smaller workshops. I would like to once again thank the entire organizing committee of the ISIT'19. I hope they are proud of their work and accomplishments, and if they are not, I hope they do not feel too bad about that. I also would like to remind our members that this technical society is run by elected volunteers, rather than hired experts or career politicians. We all strive to do our best on a

(continued on page 24)

From the Editor

Salim El Rouayheb

I hope you have enjoyed the summer, in particular our flagship conference, the 2019 IEEE International Symposium on Information Theory (ISIT), which was held in Paris this year. We start this fall issue with the awards given to members of our society. Congratulations to all the award winners. This issue features an article by Erdal Arıkan, the 2018 Shannon awardee, titled "From Sequential Decoding to Channel Polarization and Back Again", which is a written and extended version of his Shannon Lecture delivered at ISIT. Gireeja Ranade and Christina Lee Yu, the new officers of Women in The Information Theory Society (WITHITS), update us on the latest activities within WHITHITS. We also have reports from the 11th Asia-Europe workshop (AEW11) on "Concepts in Information Theory and Communications", the 40th Symposium on Information Theory in the Benelux and The Fifth London Symposium on Information Theory (LSIT). With sadness, we conclude this issue with tributes to Robert J. McEliece and Elwyn Berlekamp who have recently passed away.

As a reminder, Announcements, news, and events intended for both the printed newsletter and the website, such as award announcements, calls for nominations, and upcoming conferences, can be submitted at the IT Society website http://www.itsoc.org. Articles and columns can be e-mailed to me at salim.elrouayheb@rutgers.edu with a subject line that includes the words "IT newsletter."

The next few deadlines are:

Oct 10, 2019 for the issue of December 2019.

Jan 10, 2020 for the issue of March 2020.

April 10, 2020 for the issue of May 2020.

Please submit plain text, LaTeX, or Word source files; do not worry about fonts or layout as this will be taken care of by IEEE layout specialists. Electronic photos and graphics should be in high resolution and sent as separate files.

Salim El Rouayheb

IEEE Information Theory Society Newsletter

IEEE Information Theory Society Newsletter (USPS 360-350) is published quarterly by the Information Theory Society of the Institute of Electrical and Electronics Engineers, Inc.

Headquarters: 3 Park Avenue, 17th Floor, New York, NY 10016-5997.

Cost is $1.00 per member per year (included in Society fee) for each member of the Information Theory Society. Printed in the U.S.A. Periodicals postage paid at New York, NY and at additional mailing offices.

Postmaster: Send address changes to IEEE Information Theory Society Newsletter, IEEE, 445 Hoes Lane, Piscataway, NJ 08854.

Table of Contents

Awards

Congratulations to the members of our community that have recently received recognition for their exceptional scholarly contributions.

CHARLES BENNETT: The 2020 Claude E. Shannon Award

The Claude E. Shannon Award is the highest honor from the IEEE Information Theory Society. The award has been instituted to honor consistent and profound contributions to the field of information theory.

JOACHIM HAGENAUER: 2019 Aaron D. Wyner Distinguished Service Award

The Aaron D. Wyner Distinguished Service Award of the IT Society has been instituted to honor an individual who has shown outstanding leadership in, and provided long-standing, exceptional service to, the Information Theory community.

Information Theory Society Paper Award

The purpose of the Information Theory Paper Award is to recognize exceptional publications in the field and to stimulate interest in and encourage contributions to fields of interest of the Society.

The 2019 award winning publication is:

- E. Candes, X. Li, M. Soltanolkotabi, "Phase Retrieval via Wirtinger Flow: Theory and Algorithms", IEEE Transactions on Information Theory, Apr. 2015.

2018 IEEE Communications Society & Information Theory Society Joint Paper Award

Recognizes the author(s) of outstanding papers appearing in any publication of the IEEE Communications Society or the IEEE Information Theory Society in the previous three calendar years.

Arash Gholami Davoodi and Syed Ali Jafar
"Aligned Image Sets Under Channel Uncertainty: Settling Conjectures on the Collapse of Degrees of Freedom Under Finite Precision CSIT," *IEEE Transactions on Information Theory*, Volume 62, No. 10, pp. 5603–5618, October 2016.

2019 IEEE Communications Society & Information Theory Society Joint Paper Award

Yuyi Mao, Jun Zhang, and Khaled B. Letaief
"Dynamic Computation Offloading for Mobile-Edge Computing With Energy Harvesting Devices," *IEEE Journal on Selected Areas in Communications*, Volume 34, No. 12, pp. 3590–3605, December 2016.

SALMAN AVESTIMEHR: 2019 James L. Massey

Research & Teaching Award for young scholars recognizes outstanding achievement in research and teaching by young scholars in the Information Theory community.

DAVID SUTTER: 2019 Thomas M. Cover Dissertation Award

The IEEE Information Theory Society Thomas M. Cover Dissertation Award, established in 2013, is awarded annually to the author of an outstanding doctoral dissertation contributing to the mathematical foundations of any of the information sciences within the purview of the Society.

- D. Sutter, "Approximate Quantum Markov Chains", SpringerBriefs in Mathematical Physics, vol 28. Springer, Cham, April 2018.

Jack Keil Wolf ISIT Student Paper Award

The IEEE Jack Keil Wolf ISIT Student Paper Award is given to up to 3 outstanding papers for which a student is the principal author and presenter. The award is based on the paper's technical contribution as well as the quality of its presentation. The prize was awarded to 2 papers this year:

- P. Pandit, M. Sahraee, S. Rangan and A. K. Fletcher, "Asymptotics of MAP Inference in Deep Networks", IEEE International Symposium on Information Theory (ISIT), Paris, France, 2019.

- J. Sima and J. Bruck, "Optimal k-Deletion Correcting Codes", IEEE International Symposium on Information Theory (ISIT), Paris, France, 2019.

2019 Chapter of the Year Award

The Chapter of the Year Award recognizes a chapter that has provided their membership with the best overall set of programs and activities. The 2018 winner is the

- Japan Section Chapter: Hiroki Koga (chair) and Yuichi Kaji (vice-chair)

KANNAN RAMCHANDRAN: 2019 Padovani Lecturer

The Padovani Lecture is held annually at the North-American School of Information Theory.

H. VINCENT POOR: Benjamin Garver Lamme Award

H. Vincent Poor has been awarded the Benjamin Garver Lamme Award, honoring more than four decades of contributions to engineering education.

The Lamme Award comes from the American Society for Engineering Education, established by a group of professors, during the 1893 Chicago World's Fair, who believed engineering education should eschew then-popular apprenticeship models to stress teaching the fundamentals of science and mathematics.

ISIT 2019 Awards Ceremony

From Sequential Decoding to Channel Polarization and Back Again

Erdal Arıkan
Department of Electrical and Electronics Engineering
Bilkent University, Ankara, 06800, Turkey

***Abstract*—This note is a written and extended version of the Shannon Lecture I gave at 2019 International Symposium on Information Theory. It gives an account of the original ideas that motivated the development of polar coding and discusses some new ideas for exploiting channel polarization more effectively in order to improve the performance of polar codes.**

I. Introduction

We begin with the usual setup for the channel coding problem, as shown in Fig. 1. A message source produces a source word $\mathbf{d} = (d_1, \ldots, d_K)$ uniformly at random over all possible source words of length K over a finite set, the source word $\mathbf{d}$ is encoded into a codeword $\mathbf{x} = (x_1, \ldots, x_N)$, the codeword $\mathbf{x}$ is transmitted over a channel, the channel produces an output word $\mathbf{y} = (y_1, \ldots, y_N)$, and a decoder processes $\mathbf{y}$ to produce an estimate $\hat{\mathbf{d}} = (\hat{d}_1, \ldots, \hat{d}_K)$ of the source word $\mathbf{d}$. The performance metrics for the system are the probability of frame error $P_e = \Pr(\hat{\mathbf{d}} \neq \mathbf{d})$, the code rate $R = K/N$, and the complexity of implementation of the encoder and decoder.

Fig. 1. Channel coding system.

Shannon [1] proved that for a broad class of channels, there exists a channel parameter C, called capacity, such that arbitrarily reliable transmission (small P_e) is attainable at any given rate R if $R < C$ (and unattainable if $R > C$). Shannon's theorem settled the question about the trade-off between the rate (R) and reliability (P_e) in a communication system. However, the random-coding analysis Shannon used to prove the attainability part of his theorem left out complexity issues. Below, we present a track of ideas, as shown in Fig. 2, for constructing practically implementable codes that meet Shannon's capacity bound while providing reliable communication.

For the rest of the note, we restrict attention to binary-input memoryless channels (BMCs). By convention, the channel input alphabet will be $\{0, 1\}$, the channel output alphabet will be arbitrary, and the channel transition probabilities will be denoted by $W(y|x)$. We will also assume that the source alphabet is binary so that $\mathbf{d} \in \{0, 1\}^K$.

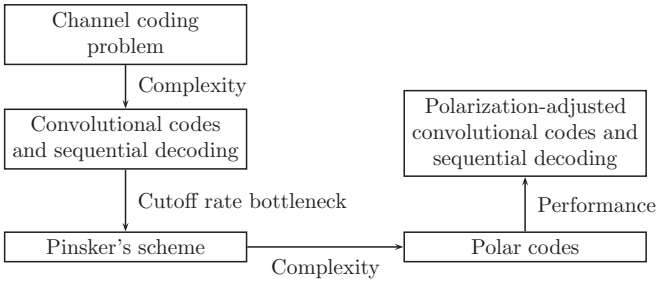


Fig. 2. Order of main topics discussed in the note.

Two channel parameters of primary interest will be the symmetric versions of channel capacity and cutoff rate, which are defined respectively as

$$C(W) = \sum_y \sum_{x \in \{0,1\}} \frac{1}{2} W(y|x) \log_2 \frac{W(y|x)}{\frac{1}{2}W(y|0) + \frac{1}{2}W(y|1)} \tag{1}$$

and

$$R_0(W) = 1 - \log_2 \left(1 + \sum_y \sqrt{W(y|0)W(y|1)}\right). \tag{2}$$

If the BMC under consideration happens to have some symmetry properties as defined in [4, p. 94], then the symmetric capacity and symmetric cutoff rate coincide with their true versions (which are obtained by an optimization over all possible distributions on the channel input alphabet). For our purposes, the symmetric versions of the capacity and cutoff rate are more relevant than their true versions since throughout this note we will be considering *linear* codes. Linear codes are constrained to use the channel input symbols 0 and 1 with equal frequency so they can at best achieve the symmetric capacity and symmetric cutoff rate. For brevity, in the rest of the note, we will omit the qualifier "symmetric" when referring to $C(W)$ and $R_0(W)$; the reader should remember that all such references are actually to the symmetric versions of these parameters as defined by (1) and (2).

A third channel parameter that will be useful in the following is the Bhattacharyya parameter defined as

$$Z(W) = \sum_y \sqrt{W(y|0)W(y|1)}. \tag{3}$$

We note the relation $R_0(W) = 1 - \log_2 \left[1 + Z(W)\right]$, which will be important in the sequel.

II. Convolutional codes and sequential decoding

Convolutional codes are a class of linear codes introduced by Elias [2] with an encoder mapping of the form $\mathbf{x} = \mathbf{dG}$ where the generator matrix $\mathbf{G}$ has a special structure that corresponds to a convolution operation. An example of a convolutional code is one with the generator matrix

$$\mathbf{G} = \begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 1 & & & & & & \\ & & 1 & 1 & 1 & 0 & 1 & 1 & & & & \\ & & & & 1 & 1 & 1 & 0 & 1 & 1 & & \\ & & & & & & 1 & 1 & 1 & 0 & 1 & 1 \end{bmatrix},$$

for which the encoding operation can be implemented using the convolution circuit in Fig. 3.

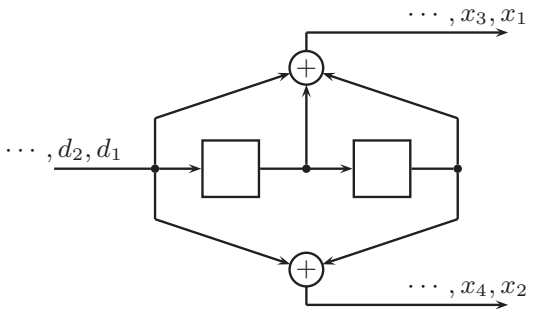


Fig. 3. Example of a convolutional code.

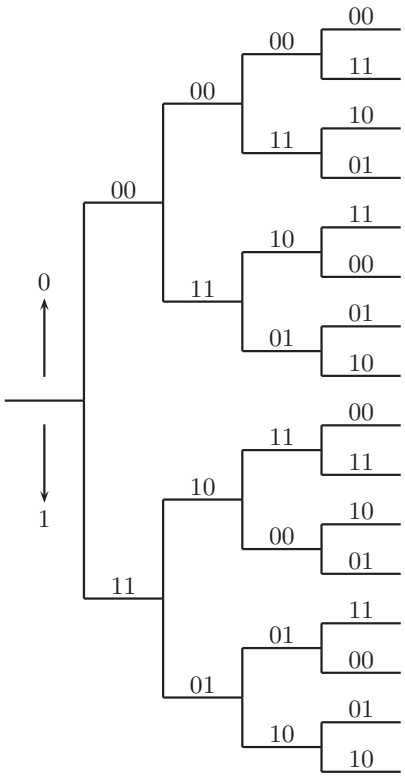


Fig. 4. Tree representation of a convolutional code.

The codewords of a convolutional code can be represented in the form of a tree. For example, the first four levels of the tree corresponding to the convolutional code of Fig. 3 are shown in Fig. 4. Each source word $\mathbf{d} = (d_1, \ldots, d_K)$ defines a path through the code tree (take the upper branch if d_i is 0, the lower branch otherwise). Branches along a path are labeled with the codeword symbols corresponding to that path.

The tree representation of a convolutional code turns the decoding problem into a tree search problem. One of the paths through the tree is the correct path and all other paths are incorrect paths. Exhaustive search of the tree for the correct path corresponds to optimum decoding but is too complex to implement. There is need for low-complexity tree search heuristics that can be used as decoders. A reasonable choice is a depth-first search heuristic. Sequential decoding is a depth-first search heuristic developed by Wozencraft [3] for decoding arbitrary tree codes.

The computational complexity in sequential decoding (the number of steps it takes to complete decoding) is a random variable whose statistical properties (mean, variance, distribution) depend on the code rate and the channel characteristics. Sequential decoding achieves the capacity $C(W)$ of any given BMC W if no limit is placed on its search complexity. However, the average complexity in sequential decoding becomes prohibitive for practical purposes if the code rate is above the cutoff rate $R_0(W)$. More precisely, at rates $R > R_0(W)$, the average complexity of decoding the first nR source bits *correctly* is lower-bounded roughly as $2^{n[R-R_0(W)]}$, while at rates $R < R_0(W)$ virtually error free communication is possible at constant average complexity per decoded bit. Detailed accounts of the sequential decoding algorithm and its complexity may be found in [4, pp. 263-286] and [5, pp. 425-476].

My interest in sequential decoding goes back to 1983 when I was a doctoral student at M.I.T. and my thesis supervisor Bob Gallager asked me to look at sequential decoding for multiaccess channels. This subject became my PhD thesis [6]. Multiaccess communications was an emerging subject and sequential decoding was a good starting point for assessing the practical viability of coding for multiaccess channels (see [7] for the broader context of this problem). Historically, sequential decoding had been a method of choice briefly (used in space communications (Pioneer 9, 1968)) before being superseded by Viterbi decoding in the 1970s. Despite having fallen out of favor, sequential decoding was still an interesting subject with rich connections to information theory and error exponents. In studying sequential decoding, I came across two fascinating papers by Pinsker [8] and Massey [9]. These papers showed how to "boost" the cutoff rate of sequential decoding in a sense described below. An extended discussion of both papers as they relate to my later work on polar coding can be found in [10]. In the following, I will focus mainly on [8] because of its general nature. However, before proceeding to [8], I will review [9] since it contains some of the essential ideas in this note in a very simple setting.

III. Massey's example

Let $M = 2^m$ for some integer $m \geq 2$, and consider an M'ary erasure channel (MEC) with input alphabet $\mathcal{X} = \{0, 1, \ldots, 2^m - 1\}$, output alphabet $\mathcal{Y} = \mathcal{X} \cup \{?\}$ (where ? is an erasure symbol), and transition probabilities $W(y|x)$ such that, when $x \in \mathcal{X}$ is sent, the channel output y has two possible values, $y = x$ and $y =?$, which it takes with conditional probabilities $W(x|x) = 1 - \epsilon$ and $W(?|x) = \epsilon$. The capacity and cutoff rate of the MEC are readily calculated as $C(m) = m(1-\epsilon)$ and $R_0(m) = m - \log_2\left(1 + (2^m - 1)\epsilon\right)$.

Massey observed that the MEC can be split into m binary erasure channels (BECs) by relabeling its inputs and outputs with vectors of length m. A specific labeling that achieves this is as follows. Each input symbol $x \in \mathcal{X}$ is relabeled with its binary representation $(x_1, \ldots, x_m) \in \{0,1\}^m$ so that $x = \sum_{i=1}^m x_i 2^{m-i}$. Each output symbol $y \in \mathcal{Y}$ is relabeled with a vector $(y_1, \ldots, y_m)$ which equals the binary representation of y if $y \in \mathcal{X}$ and equals $(?, \ldots, ?)$ if $y =?$. With this relabeling, a single transmission event $\{(x_1, \ldots, x_m) \to (y_1, \ldots, y_m)\}$ across the MEC can be thought of as a collection of m transmission events $\{x_i \to y_i\}$ across the coordinate channels. An erasure event in the MEC causes an erasure event in all coordinate channels; if there is no erasure in the MEC, there is no erasure in any of the coordinate channels. Each coordinate channel is a BEC with erasure probability ϵ. The coordinate channels are fully correlated in the sense that when an erasure occurs in one of them, an erasure occurs in all of them.

The capacity and cutoff rate of the BECs are given by $C(1) = 1 - \epsilon$ and $R_0(1) = 1 - \log_2(1 + \epsilon)$. It can be verified readily that $C(m) = mC(1)$ (capacity is conserved), while $R_0(m) \leq mR_0(1)$ with strict inequality unless ϵ equals 0 or 1. Thus, splitting the MEC does not cause a degradation in channel capacity but "improves" or "boosts" the cutoff rate. This example shows that one may break the cutoff rate barrier for the MEC by employing a separate convolutional encoder – sequential decoder pair on each coordinate BEC. The reader is advised to see [7] for an alternative look at this important example from the perspective of multiaccess channels. To learn about the communications engineering context in which Massey's example arose, we refer to [9].

Massey's example provides a basis for understanding the more complex schemes presented below. These more complex schemes begin with independent copies of a binary-input channel (raw channels), build up a large channel (akin to the MEC) through some channel combining operations, and then split the large channel back to a set of correlated binary-input channels (synthesized channels). One speaks of a "boosting" of the cutoff rate if the sum of the cutoff rates of the synthesized channels is larger than the sum of the cutoff rates of the raw channels.

IV. Pinsker's scheme

Pinsker [8] observed that, for the binary symmetric channel (BSC) with crossover probability p (a BMC with output alphabet $\{0,1\}$ and $W(1|0) = W(0|1) = p$), the ratio of the cutoff rate to capacity approaches 1 as p goes to 0,

$$\frac{R_0}{C} = \frac{1 - \log_2\left[1 + 2\sqrt{p(1-p)}\right]}{1 + p\log_2(p) + (1-p)\log_2(1-p)} \to 1 \quad \text{as } p \to 0,$$

as illustrated in Fig. 5. Pinsker combined this observation with Elias' product coding idea [11] to construct a coding scheme that boosted the cutoff rate to capacity.

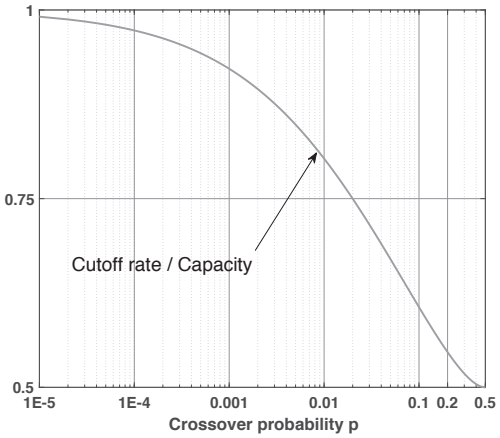


Fig. 5. Ratio of cutoff rate to capacity for the BSC.

Pinsker's scheme, as shown in Fig. 6, uses an inner block code and K identical outer convolutional codes. Each round of operation of the inner block code comprises the encoder for the inner block code receiving one bit from the output of each outer convolutional encoder (for a total of K bits) and encoding them into an inner code block of length N bits. The inner code block is then sent over a BMC W by N uses of W. Since successive bits at the output of each outer convolutional encoder are carried in separate inner code blocks, they suffer i.i.d. error events. So, each outer convolutional code sees a *memoryless* bit-channel, as depicted in Fig. 7. We denote by $W_i : U_i \to \hat{U}_i$ the (virtual) BMC that connects the ith convolutional encoder to the ith sequential decoder.[1]

To show that this scheme is capable of boosting the cutoff rate arbitrarily close to channel capacity, we may fix the rate K/N of the inner block code as $(1 - \delta)C(W)$ for some constant $0 < \delta < 1$ and consider increasing the block length N and choosing a good enough inner block code so as to ensure that the bit-channels $W_1, \ldots, W_K$ become near-perfect with $R_0(W_i) > 1 - \epsilon$ for each i, where $\epsilon > 0$ is a second constant independent of N and i. This ensures that each outer convolutional code can operate at a rate $1 - \epsilon$ and still be decoded by a sequential decoder at an average complexity bounded by a third constant, where the third constant depends on δ and ϵ but not on N. The overall rate for this scheme is $K(1-\epsilon)/N = (1-\delta)(1-\epsilon)C(W)$, which can be made

[1] We use capital letters U_i and $\hat{U}_i$ to denote the random variables corresponding to u_i and $\hat{u}_i$. This convention of using capital letters to denote random variables is followed throughout.

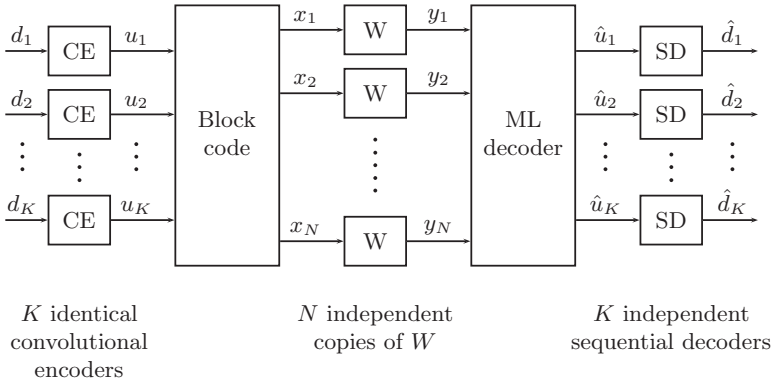


Fig. 6. Pinsker's scheme.

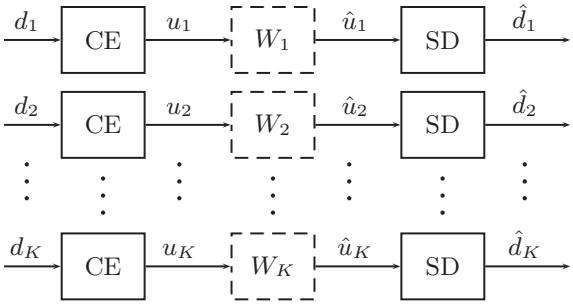


Fig. 7. Bit-channels created by Pinsker's scheme.

arbitrarily close to $C(W)$ by choosing δ and ϵ sufficiently small. In Pinsker's words, his scheme shows that "[f]or a very general class of channels operating below capacity it is possible to construct a code in such a way that the number of operations required for decoding is less than some constant that is independent of the error probability".

Pinsker's result complements Shannon's result by showing that, at any fixed rate R below channel capacity $C(W)$, the average complexity per decoded bit can be kept bounded by a constant while achieving any desired frame error rate $P_e > 0$. Unfortunately, the recipe for choosing a good enough inner block code in Pinsker's scheme is to pick the code at random. The non-constructive nature of Pinsker's scheme and the complexity of ML decoding of a randomly chosen block code make Pinsker's scheme impractical. For our purposes, the takeaway from Pinsker's scheme is the demonstration that there is no "cutoff rate barrier to sequential decoding" in a fundamental sense. Our next goal will be to find a way of breaking the cutoff rate barrier in a practically implementable manner.

Before we end this section, it is instructive to compare Pinsker's scheme with Massey's example. In Massey's example, a given channel is split into multiple correlated bit-channels. In Pinsker's scheme, the first step is to synthesize a large channel from a collection of independent bit-channels; the large channel is then split back into a number of dependent bit-channels. Massey's example appears to be a very special case that cannot be generalized to arbitrary BMCs, while Pinsker's scheme is entirely general. Massey's example boosts the cutoff rate almost effortlessly but cannot boost it all the way to channel capacity. Pinsker's scheme is much more complex but can boost the cutoff rate to capacity. Both schemes use multiple sequential decoders. The use of multiple sequential decoders is a crucial aspect of both schemes. If a single sequential decoder were used in Pinsker's scheme to decode all K convolutional codes jointly (using a joint tree representation), then a "data-processing" theorem by Gallager [4, pp. 149-150] would limit the achievable cutoff rate to $R_0(W)$. For more on this point, we refer to [10].

V. Multi-level coding

In order to reduce the complexity in Pinsker's scheme, in this section, we look at multi-level coding (MLC) with multi-stage decoding (MSD), a scheme due to Imai and Hirakawa [12]. The MLC/MSD system makes better use of the information available at the receiver and hence it has the potential to boost the cutoff rate at lower complexity. The particular MLC/MSD system we consider here is shown in Fig. 8. The mapper in the figure is a one-to-one transformation. The demapper is a device that calculates sufficient statistics in the form of log-likelihood ratios (LLRs) and feeds them to a MSD unit. Each decoder in the MSD chain is able to benefit from the decisions by the previous decoders in the chain.

In effect, the MLC/MSD system creates N bit-channels $W_1, \ldots, W_N$, as shown in Fig. 9, where the ith bit-channel is of the form $W_i : U_i \to \mathbf{Y}\hat{\mathbf{U}}^{i-1}$. More precisely, W_i is the channel whose input U_i is a bit taken from the output of the ith convolutional encoder and whose output $\mathbf{Y}\hat{\mathbf{U}}^{i-1}$ is the input to the ith sequential decoder in the MSD chain. Here, $\mathbf{Y} = (Y_1, \ldots, Y_N)$ is the entire channel output vector and $\hat{\mathbf{U}}^{i-1} = (\hat{U}_1, \ldots, \hat{U}_{i-1})$ is the vector of decisions provided by the decoders that precede decoder i in the MSD chain.

If the MLC/MSD system is configured so that the sequential decoders provide virtually error-free decisions, then the bit-channel W_i takes the form $W_i : U_i \to \mathbf{Y}\mathbf{U}^{i-1}$ where the decisions fed forward by the previous stages are always

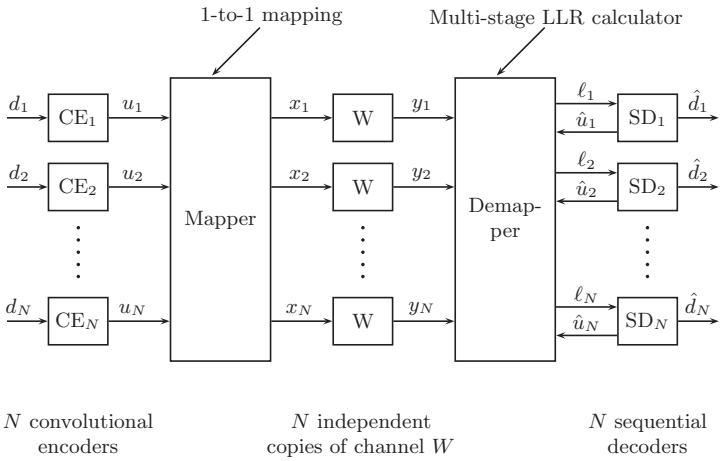


Fig. 8. Multi-level coding

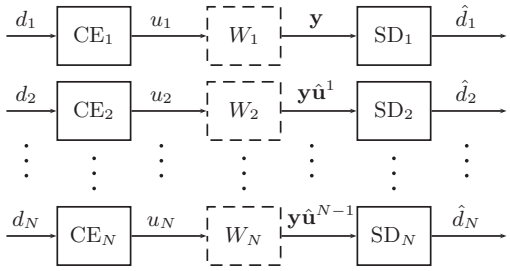


Fig. 9. Bit channels created by MLC/MSD

correct. For purposes of deriving polar codes, it suffices to consider only this ideal case with no decision errors. Hence, from now on, we suppose that W_i has this ideal form.

An important property of the MLC/MSD scheme is the conservation of capacity,

$$\sum_{i=1}^{N} C(W_i) = \sum_{i=1}^{N} I(U_i; \mathbf{Y}\mathbf{U}^{i-1}) = I(\mathbf{U}^N; \mathbf{Y}^N) = NC(W),$$

where the second equality is obtained by writing $I(U_i; \mathbf{Y}\mathbf{U}^{i-1}) = I(U_i; \mathbf{Y}|\mathbf{U}^{i-1})$ based on the assumption that U_i and $\mathbf{U}^{i-1}$ are independent and then using the chain rule.

The MLC/MSD scheme conserves capacity at any finite construction size N while Pinsker's scheme conserves capacity only in an asymptotic sense. Thus MLC/MSD uses information more efficiently and hence may be expected to achieve a given performance at a lower construction size (leading to a lower complexity).

On the other hand, unlike Pinsker's scheme in which the outer convolutional codes are all identical, the natural rate assignment for the MLC/MSD scheme is to set the rate R_i of the ith convolutional code to a value just below $R_0(W_i)$. Using convolutional codes at various different rates $\{R_i\}$ as dictated by $\{R_0(W_i)\}$, and decoding them using a chain of sequential decoders is a high price to pay for the greater information efficiency of the MLC/MSD scheme. Fortunately, this complexity issue regarding outer convolutional codes and sequential decoders is not as severe as it looks thanks to a phenomenon called *channel polarization*.

Theorem 1: Consider a sequence of MLC/MSD schemes over a BMC W, with the nth scheme in the sequence having size $N = 2^n$ and a mapper of the form

$$\mathbf{P}_n = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}^{\otimes n}, \tag{4}$$

where the exponent "$\otimes n$" indicates the nth Kronecker power. Fix $0 < \delta < \frac{1}{2}$. As n increases, the idealized bit-channels $\{W_i\}_{i=1}^N$ for the nth MLC/MSD scheme polarize in the sense that the fraction of channels with $C(W_i) > 1 - \delta$ tends to $C(W)$ and the fraction with $C(W_i) < \delta$ tends to $1 - C(W)$. For each bit-channel W_i that polarizes, its cutoff rate $R_o(W_i)$ polarizes to the same point (0 or 1) as its capacity $C(W_i)$. Furthermore, the mapper and demapper functions can be implemented at complexity $\mathcal{O}(N \log N)$ per mapper block $\mathbf{u}$. $\diamond$

We refer to [13] for a proof of this theorem.

The most important aspect of Theorem 1 is its statement that polarization can be achieved at complexity $\mathcal{O}(\log N)$ per transmitted bit. In the absence of a complexity constraint, polarization alone is not hard to achieve. A randomly chosen mapper is likely to achieve polarization but is also likely to be too complex to implement. The recursive structure of the mappers $\{\mathbf{P}_n\}$ used in Theorem 1 make it possible to obtain polarization at low complexity. We will see below that the polarization effect brought about by the transforms $\{\mathbf{P}_n\}$ is strong enough to simplify the rate assignment $\{R_i\}$ while also maintaining reliable transmission of source data bits after the MLC/MSD scheme is simplified. However, we first wish to illustrate the polarization phenomenon of Theorem 1 by an example.

In Fig. 10, we show a plot of $C(W_i)$ v. i for the bit-channels $\{W_i\}$ created by an MLC/MSD construction of size $N = 128$ using the transform $\mathbf{P}_n$ with $n = 7$. The channel in

the example is a binary-input additive white Gaussian noise (BIAWGN) channel, which is a channel that receives a binary symbol $x \in \{0,1\}$ as input, maps it into a real number s by setting $s = 1$ if $x = 0$ and $s = -1$ otherwise, and generates a channel output $y = s + z$, where $z \sim N(0, \sigma^2)$ is additive Gaussian noise independent of s. The signal-to-noise ratio (SNR) for the BIAWGN channel is defined as $1/\sigma^2$. The SNR in Fig. 10 is 3 dB. The capacity $C(W)$ of the BIAWGN channel W at 3 dB SNR is 0.72 bits; hence, by Theorem 1, we expect that roughly a fraction 0.72 of the capacity terms $C(W_i)$ in Fig. 10 will be near 1.

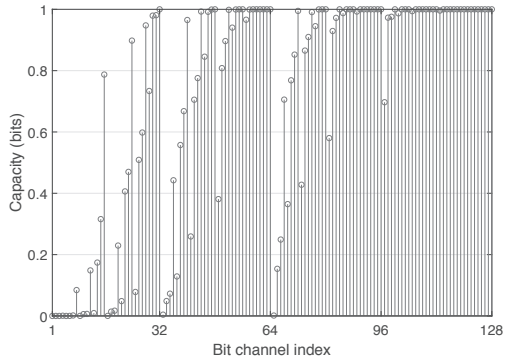


Fig. 10. Channel polarization for BIAWGN channel at 3 dB SNR.

An alternative view of the channel polarization effect in the preceding example is presented in Fig. 11 where cumulative distributions (*profiles*) of various information parameters are plotted as a function of an index variable i which takes values from 0 to $N = 128$. The polarized capacity profile is defined as the sequence of cumulatives $\left\{\sum_{j=1}^{i} C(W_j)\right\}$ indexed by i. Likewise, the polarized cutoff rate profile is defined as $\left\{\sum_{j=1}^{i} R_0(W_j)\right\}$, the unpolarized capacity profile as $\left\{iC(W)\right\}$, and the unpolarized cutoff rate profile as $\left\{iR_0(W)\right\}$. By convention, we start each profile at 0 at $i = 0$. The two other curves in the figure (Reed-Muller and polar code rate profiles) will be discussed later.

The unpolarized capacity and cutoff rate profiles in Fig. 11 serve as benchmarks, corresponding to the case where the mapper in the MLC scheme is the identity transform. The polarized capacity and cutoff rate profiles demonstrate the polarization effect due to the transform $\mathbf{P}_7$. The polarized and unpolarized capacity profiles coincide at $i = 0$ and $i = N$, but a gap exists between the two for $0 < i < N$ due to channel polarization. Ideally, the polarized capacity profile would stay zero until i is around $[1 - C(W)]N = 35.8$ and then climb with a slope of 1 until $i = N$. A mapper chosen at random is likely to create a near-ideal polarized capacity profile, but the corresponding demapper function is also likely to be too complex. By using $\mathbf{P}_7$ as the mapper, we settle for a non-ideal polarized capacity profile in return for lower implementation complexity.

A beneficial by-product of channel polarization is the boosting of the cutoff rate, which is clearly visible in Fig. 11. The

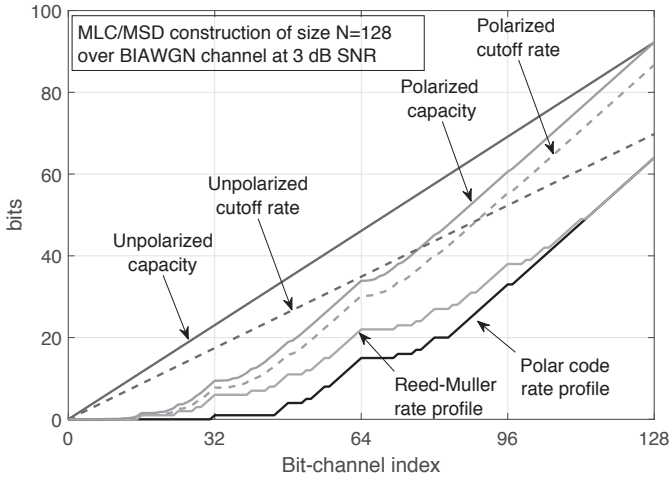


Fig. 11. Capacity and cutoff rate profiles over BIAWGN channel.

polarized cutoff rate profile has a final value $\sum_{i=1}^{N} R_0(W_i) = 86.7$ compared to a final value $NR_0(W) = 69.8$ for the unpolarized cutoff rate profile. Theorem 1 ensures that, asymptotically as N becomes large, the normalized sum cutoff rate $\frac{1}{N}\sum_{i=1}^{N} R_0(W_i)$ approaches $C(W)$. So, the MLC/MSD scheme, equipped with the transforms $\{\mathbf{P}_n\}$, reproduces Pinsker's result by boosting the cutoff rate to channel capacity, with the important difference that here the mapper and demapper complexity per transmitted source bit is $\mathcal{O}(\log N)$ for a construction of size N (while the similar complexity in Pinsker's scheme is exponential in N).

Despite the reduced mapper/demapper complexity, the MLC/MSD scheme (with the transforms $\{\mathbf{P}_n\}$) is still far from being practical since it calls for using N outer convolutional codes at various code rates. At this point, we take advantage of the polarization effect and constrain the rates R_i to 0 or 1. Such a 0-1 rate assignment in effect eliminates the outer codes. Setting $R_i = 0$ corresponds to fixing the input to the ith bit channel W_i. Setting $R_i = 1$ corresponds to sending information in uncoded form over the ith bit-channel W_i. In either case, the MSD decisions can be made independently from one mapper block (of length N) to the next, eliminating the need for a sequential decoder.

The 0-1 rate assignment leads to a new type of stand-alone block code, which we will call a *polar code*. The simplified MSD function under the 0-1 rate assignment will be called *successive cancellation* (SC) decoding. An important new question that arises is whether polar codes, obtained by such drastic simplification of the MLC/MSD scheme, can provide reliable transmission of source data. An answer to this question is provided in the next section.

VI. Polar Codes

In this section we will study polar codes as a stand-alone coding scheme. For simplicity, we will consider polar coding only for BMCs that are symmetric in the sense defined in [13]

or [4, p. 94]. We begin by restating the definition of polar codes without any reference to their origin.

A *polar code* is a linear block code characterized by three parameters: a code block-length N, a code dimension K, and a data index set $\mathcal{A}$. The code block-length is constrained to be a power of two, $N = 2^n$ for some $n \geq 1$. The code dimension can be any integer in the range $1 \leq K \leq N$. The data index set $\mathcal{A}$ is a subset of $\{1, \ldots, N\}$ with size $|\mathcal{A}| = K$. (This set corresponds to the set of indices i for which $R_i = 1$ in the MLC/MSD context.) A method of choosing $\mathcal{A}$ will be given below. The encoder for a polar code with parameters $(N, K, \mathcal{A})$ receives a source word $\mathbf{d}$ of length K and embeds it in a carrier vector $\mathbf{u}$ so that $\mathbf{u}_{\mathcal{A}} = \mathbf{d}$ and $\mathbf{u}_{\mathcal{A}^c} = \mathbf{0}$. (Here, $\mathbf{u}_{\mathcal{A}} = (u_i : i \in \mathcal{A})$ is a subvector of $\mathbf{u}$ obtained by discarding all coordinates outside $\mathcal{A}$.) Encoding is completed by computing the transform $\mathbf{x} = \mathbf{u}\mathbf{P}_n$, where $\mathbf{P}_n$ is as defined in (4). Henceforth, we will refer to $\mathbf{P}_n$ as a *polar transform*.

The standard decoding method for polar codes is SC decoding. For details of SC decoding, we refer to [13]. As shown in [13], for a symmetric BMC W, the probability of frame error P_e for a polar code under SC decoding is bounded as

$$P_e \leq \sum_{i \in \mathcal{A}} Z(W_i) \tag{5}$$

where $Z(W_i)$ is the Bhattacharyya parameter of channel W_i. From now on, we will assume that the data index set $\mathcal{A}$ is chosen so as to minimize the bound (5) on P_e, *i.e.*, that $\mathcal{A}$ is selected as a set of K indices i such that $Z(W_i)$ is among the K smallest numbers in the list $Z(W_1), \ldots, Z(W_N)$. Since $Z(W_i) = 2^{1-R_0(W_i)} - 1$, an equivalent rule for constructing a polar code is to select $\mathcal{A}$ as a set of K indices i such that $R_0(W_i)$ is among the K largest cutoff rates in the list $R_0(W_1), \ldots, R_0(W_N)$.

Theorem 2: A polar code with length N, dimension K, and rate $R = K/N$ over a symmetric BMC W has the following properties.

- It can be constructed (the data index set $\mathcal{A}$ can be determined) in $\mathcal{O}(N\text{poly}(\log N))$ steps [14], [15], [16].
- It can be encoded and SC-decoded in $\mathcal{O}(N \log N)$ steps [13].
- Its frame error rate P_e under SC decoding is bounded as $\mathcal{O}(e^{-N^{0.499}})$ for any fixed rate $R < C(W)$ [17].

$\diamond$

In summary, polar coding achieves the capacity of symmetric BMCs with low-complexity encoding, decoding, and construction methods. For a precise discussion of the novelty of polar codes as a capacity-achieving code construction, we refer to [18].

The performance of polar codes is far from optimal. Fig. 12 illustrates the frame error rate (FER) P_e under SC decoding of a polar code with block-length $N = 128$ and rate $R = 1/2$ over a BIAWGN channel with the SNR ranging from 0 to 5 dB. This and other FER curves in Fig. 12 have been obtained by computer simulation. Also shown in Fig. 12 is the BIAWGN dispersion approximation [19] at block-length $N = 128$ and rate $R = 1/2$, which is an estimate of the average ML-decoding performance over the BIAWGN channel of a code chosen uniformly at random from the ensemble of all possible binary codes of block-length $N = 128$ and rate $R = 1/2$.

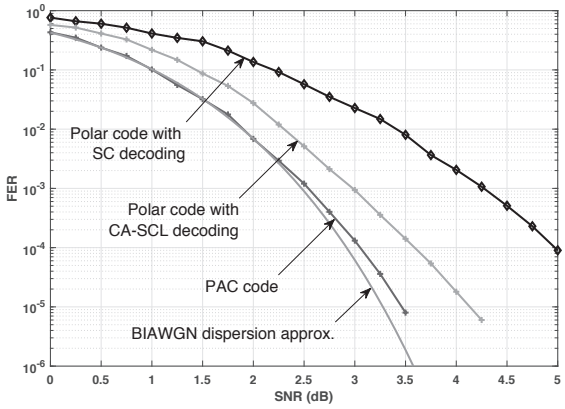


Fig. 12. Performance curves over the BIAWGN channel.

The weak performance of polar codes is due in part to the suboptimality of the SC decoder and in part to the poor minimum distance of polar codes. An effective method to fix both of these problems has been to use a concatenation scheme in which a high-rate outer code is used to pre-code the source bits before they go into an inner polar code. A particularly powerful example of such methods is the CRC-aided SC list decoding (CA-SCL) [20], whose FER performance is shown in Fig. 12 for the case of $N = 128$, $R = 1/2$, CRC length 8, and list size 32. In the next section, we consider improving the polar code performance still further by shifting the burden of error correction entirely to an outer code.

VII. Polarization-adjusted convolutional codes

In this section, we consider a new class of codes that we will refer to as *polarization-adjusted convolutional* (PAC) codes. The motivating idea for PAC codes is the recognition that 0-1 rate assignments waste the capacities $C(W_i)$ of bit-channels W_i whose inputs are fixed by the rate assignment $R_i = 0$. The capacity loss is especially significant at practical (small to moderate) block-lengths N since polarization takes place relatively slowly. In order to prevent such capacity loss, we need a scheme that avoids fixing the input of any bit-channel. PAC codes achieve this by placing an outer convolutional coding block in front of the polar transform as shown in Fig. 13.

As with polar codes, the natural block lengths for PAC codes are powers of two, $N = 2^n$, $n \geq 1$. The code dimension K can be any integer between 1 and N. The encoding operation for PAC codes is as follows. A rate-profiling block inserts the source word $\mathbf{d}$ into a data carrier word $\mathbf{v}$ in accordance with a data index set $\mathcal{A}$ so that $\mathbf{v}_{\mathcal{A}} = \mathbf{d}$ and $\mathbf{v}_{\mathcal{A}^c} = \mathbf{0}$. The PAC codeword $\mathbf{x}$ is obtained from $\mathbf{v}$ by a one-to-one transformation $\mathbf{x} = \mathbf{v}\mathbf{T}\mathbf{P}_n$ where $\mathbf{T}$ is a convolution operation and $\mathbf{P}_n$ is the

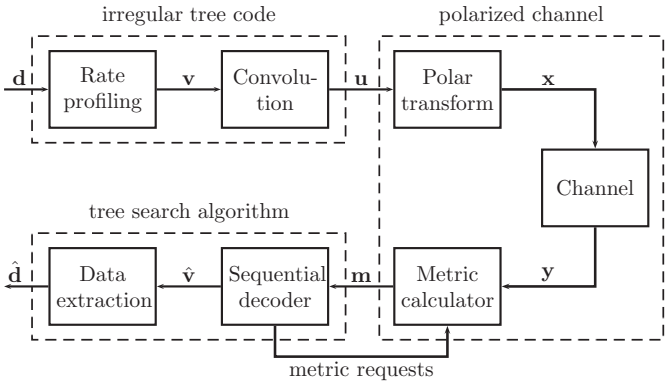


Fig. 13. PAC coding scheme.

polar transform. A low-complexity encoding alternative is to compute first $\mathbf{u} = \mathbf{vT}$ and then $\mathbf{x} = \mathbf{uP}_n$.

As usual, we characterize the convolution operation by an impulse response $\mathbf{c} = (c_0, \cdots, c_m)$, where by convention we assume that $c_0 \neq 0$ and $c_m \neq 0$. The parameter $m+1$ is called the constraint length of the convolution. The input-output relation for a convolution with a given impulse response $\mathbf{c} = (c_0, \cdots, c_m)$ is

$$u_i = \sum_{j=0}^{m} c_j v_{i-j}$$

where it is understood that $v_{i-j} = 0$ for $j \geq i$. The same convolution operation can be represented in matrix form as $\mathbf{u} = \mathbf{vT}$ where $\mathbf{T}$ is an upper-triangular Toeplitz matrix,

$$\mathbf{T} = \begin{bmatrix} c_0 & c_1 & c_2 & \cdots & c_m & 0 & \cdots & 0 \\ 0 & c_0 & c_1 & c_2 & \cdots & c_m & & \vdots \\ 0 & 0 & c_0 & c_1 & \ddots & \cdots & c_m & \vdots \\ \vdots & 0 & \ddots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \ddots & 0 & \vdots \\ \vdots & & & \ddots & 0 & c_0 & c_1 & c_2 \\ \vdots & & & & 0 & 0 & c_0 & c_1 \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & 0 & c_0 \end{bmatrix}.$$

To illustrate the above encoding operation, consider a small example with $N = 8$, $K = 4$, $\mathcal{A} = \{4, 6, 7, 8\}$, and $\mathbf{c} = (1, 1, 1)$. The rate-profiler maps the source word $\mathbf{d} = (d_1, \ldots, d_4)$ into $\mathbf{v} = (v_1, \ldots, v_8)$ so that

$$\mathbf{v} = (0, 0, 0, d_1, 0, d_2, d_3, d_4).$$

The convolution $\mathbf{u} = \mathbf{vT}$ generates an output word $\mathbf{u}$ with $u_1 = v_1$, $u_2 = v_1 + v_2$, and $u_i = v_{i-2} + v_{i-1} + v_i$ for $i = 3, \ldots, 8$. (This convolution can be implemented as in Fig. 3 by taking the upper part of that circuit.) Encoding is completed by computing the polar transform $\mathbf{x} = \mathbf{uP}_3$.

Unlike ordinary convolutional codes, the convolution operation here generates an irregular tree code due to the constraint $\mathbf{v}_{\mathcal{A}^c} = \mathbf{0}$. Fig. 14 illustrates the irregular tree code generated by the convolution in the above example. The tree in Fig. 14 branches only at time indices in the set $\mathcal{A}$, *i.e.*, only when there is a new source bit d_i going into the convolution operation. When there is a branching in the tree at some stage $i \in \mathcal{A}$, by convention, the upper branch corresponds to $v_i = 0$ and the lower branch to $v_i = 1$. Leaf nodes of the tree in Fig. 14 are in one-to-one correspondence with the convolution input words $\mathbf{v}$ satisfying the constraint $\mathbf{v}_{\mathcal{A}^c} = \mathbf{0}$. The branches on the path to a leaf node $\mathbf{v}$ are labeled with the symbols of the convolution output word $\mathbf{u} = \mathbf{vT}$.

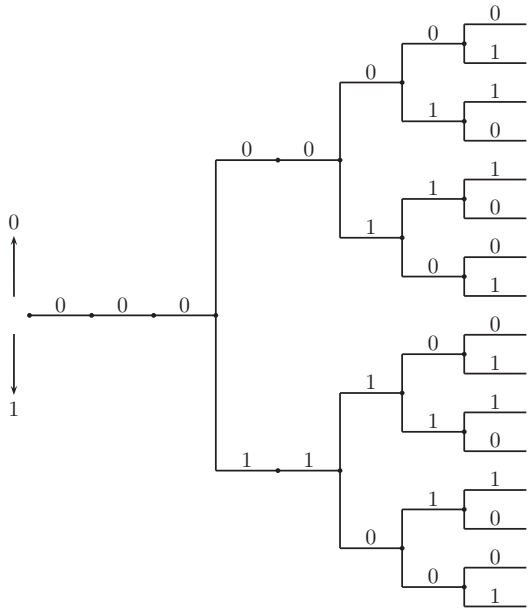


Fig. 14. Irregular tree code example.

To summarize, a PAC code is specified by four parameters $(N, K, \mathcal{A}, \mathbf{c})$. In simulation studies we observed that the performance of a PAC code is more sensitive to the choice of $\mathcal{A}$ than to $\mathbf{c}$. As long as the constraint length of the convolution is sufficiently large, choosing $\mathbf{c}$ at random may be an acceptable design practice. Finding good design rules for $\mathcal{A}$ is a research problem.

A heuristic method of choosing $\mathcal{A}$ is to use a *score* function $s : \{1, \ldots, N\} \to \mathbb{R}$ and select $\mathcal{A}$ as a set of indices i such that $s(i)$ is among the largest K scores in the list $s(1), \ldots, s(N)$ (with ties broken arbitrarily). Two examples of score functions (inspired by polar codes) are the capacity score function $s(i) = C(W_i)$ and the cutoff rate score function $s(i) = R_0(W_i)$ where $\{W_i\}$ are the MLC/MSD bit-channels created by the polar transform $\mathbf{P}_n$. The cutoff rate score function recovers polar codes when $\mathbf{T}$ is set to the identity transform (corresponding to $\mathbf{c} = 1$). A third example of a score function is the Reed-Muller (RM) score function $s(i) = w(i-1)$ where $w(i-1)$ is the number of ones in the binary representation of $i-1, 0 \leq i-1 \leq N-1$. For example, $w(12) = 2$ since 12 has the binary representation 1100. We refer to this score function as the RM score function since it

generates the well-known RM codes [22], [23] when $\mathbf{T}$ is the identity transform.

We now turn to decoding of PAC codes. For purposes of discussing the decoding operation, it is preferable to segment the PAC coding system into three functional blocks as shown by dashed-rectangles in Fig. 13. According to this functional segmentation, a source word $\mathbf{d}$ is inserted into a data carrier $\mathbf{v}$, the data carrier $\mathbf{v}$ is encoded into an codeword $\mathbf{u}$ from an irregular tree code, the codeword $\mathbf{u}$ is sent over a polarized channel, a sequential decoder is used to generate an estimate $\hat{\mathbf{v}}$ of $\mathbf{v}$, and finally, an estimate $\hat{\mathbf{d}}$ of the source word $\mathbf{d}$ is extracted from $\hat{\mathbf{v}}$ by setting $\hat{\mathbf{d}} = \hat{\mathbf{v}}_{\mathcal{A}}$.

Irregular tree codes can be decoded by tree search heuristics in much the same way as regular tree codes. A particularly suitable tree search heuristic for PAC codes is sequential decoding, specifically, the Fano decoder [21]. The Fano decoder tries to identify the correct path in the code tree by using a metric that tends to drift up along the correct path and drift down as soon as a path diverges from the correct path. The Fano decoder generates metric requests along the path that it is currently exploring and a metric calculator responds by sending back the requested metric values (denoted by $\mathbf{m}$ in Fig. 13). Unlike the usual metric in sequential decoding, the metrics here have to have a time-varying bias so as to maintain the desired drift properties in the face of the irregular nature of the tree code. In computing the metric, the metric calculator can use a recursive method, as in SC decoding of polar codes.

Fig. 12 presents the result of a computer simulation with a PAC code with $N = 128$, $R = 1/2$, $\mathcal{A}$ chosen in accordance with the RM design rule, and $\mathbf{c} = (1, 0, 1, 1, 0, 1, 1)$. As seen in the figure, the FER performance of the PAC code in this example comes very close to the dispersion approximation for FER values larger than 10^{-3}. Evidently, the product of the polar transform $\mathbf{P}_n$ and the convolution transform $\mathbf{T}$ creates an overall transform $\mathbf{G} = \mathbf{T}\mathbf{P}_n$ that looks sufficiently random to achieve a performance near the dispersion approximation. When we repeated this simulation experiment with a PAC code designed by the polar coding score function (keeping everything else the same), we observed that the performance became worse but the sequential decoder ran significantly faster. The RM design was the best design we could find in terms of FER performance.

As a heuristic guide to understanding the computational behavior of sequential decoding of a PAC code, we found it useful to associate a *rate profile* to each design rule or equivalently data index set $\mathcal{A}$. The rate profile for a data index set $\mathcal{A}$ is defined as the the sequence of numbers $\{K_i\}_{i=0}^{N}$ where $K_0 = 0$ and K_i is the number of elements in $\mathcal{A} \cap \{1, 2, \ldots, i\}$ for $i \geq 1$. Thus, K_i is the number of source data bits carried in the first i coordinates of the data carrier word $\mathbf{v}$. The rate profiles associated with the RM and polar code design rules are shown in Fig. 11 for $N = 128$ and $K = 64$. We expect that a design rule whose rate profile stays below the polarized cutoff rate profile at a certain SNR will generate a PAC code that has low complexity under sequential decoding at that SNR. In Fig. 11, both the RM and polar rate profiles lie below the polarized cutoff rate profile, but the polar rate profile leaves a greater safety margin, which may explain the experimental observation that the Fano decoder runs faster with the polar code design rule.

VIII. Remarks and Open Problems

We conclude the note with some complementary remarks about PAC codes and suggestions for further research.

One may view PAC codes as a concatenation scheme with an outer convolutional code and an inner polar code. However, PAC codes differ from typical concatenated coding schemes in that the inner code in PAC coding has rate one, so it has no error correction capability. It is more appropriate to view the inner polar transform and the metric calculator (mapper and demapper) in PAC coding as a pair of pre- and post-processing devices around a memoryless channel that provide polarized information to an outer decoder so as to increase the performance of the outer coding system.

In view of the data-processing theorem mentioned in connection with Pinsker's scheme, it seems impossible that PAC codes be able to operate at low-complexity at rates above the cutoff rate $R_0(W)$ using only a *single* sequential decoder. This is true only in part. PAC codes use a convolutional code whose length spans only one use of the polarized channel. The sequential decoder in PAC coding stops searching for the correct path if a decision error is made after reaching level N in the irregular code tree, *i.e.*, after a single use of the polarized channel. The $R_0(W)$ bound on sequential decoding would hold if a convolutional code were used that extended over multiple uses of the polarized channel. A better understanding of the computational complexity of the sequential decoder in PAC coding is an open problem.

As stated above, the performance and complexity of PAC codes are yet to be studied rigorously. It is clear that in general PAC codes can achieve channel capacity since they contain polar codes as a special case. The main question is to characterize the best attainable performance by PAC codes over variation of the data index set $\mathcal{A}$ and the convolution impulse response $\mathbf{c}$.

The fact that PAC codes perform well under the RM design rule suggests that, unlike polar codes, PAC codes are robust against channel parameter variations and modeling errors. It is of interest to investigate if PAC codes have *universal* design rules so that a given PAC code performs well uniformly over the class of all BMCs with a given capacity. In particular, it is of interest to check if the RM design rule (together with a suitably chosen convolution impulse response $\mathbf{c}$) is universal in this sense.

A disadvantage of the sequential decoding method is its variable complexity. It is of interest to study fixed-complexity search heuristics for decoding PAC codes. One possibility is to use a breadth-first search heuristic, such as a Viterbi decoder. However, a Viterbi decoder that tracks only the state of the convolutional encoder will be suboptimal since PAC codes incorporate a polarized channel that, too, has a state. In fact, the number of states of the polarized channel is the same as

the number of possible words $\mathbf{u}$ at the input of the polarized channel, namely, 2^{NR} for a PAC code of length N and rate R. There is clearly need for a sub-optimal breadth-first search heuristic that tracks only a subset of all possible states. One option that may be considered here is list Viterbi decoding [24] which is a method that has proven effective for searching large state spaces. For some other alternatives of forward pruning methods in breadth-first search, such as beam search, we refer to [25, pp. 174-175].

In linear algebra, lower-upper decomposition (LUD) is a method for solving systems of linear equations. PAC coding may be regarded as one form of upper-lower decomposition (ULD) of a code generator matrix G for purposes of solving a redundant set of linear equations when the equations are corrupted by noise. One may investigate if there are other decompositions in linear algebra for synthesizing generator matrices that yield powerful codes with low-complexity encoding and decoding.

References

[1] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, pp. 379–423, July 1948.

[2] Peter Elias, "Coding for noisy channels," in *IRE Convention Record, Part 4*, pp. 37–46, Mar. 1955.

[3] J. M. Wozencraft, "Sequential Decoding for Reliable Communication," Tech. Report 325, Res. Lab. Elect., M.I.T., Aug. 1957.

[4] R. G. Gallager, *Information Theory and Reliable Communication*. New York: Wiley, 1968.

[5] J. M. Wozencraft and I. M. Jacobs, *Principles of Communication Engineering*. New York: Wiley, 1965.

[6] E. Arıkan, "Sequential Decoding for Multiple Access Channels," Tech. Rep. LIDS-TH-1517, Lab. Inf. Dec. Syst., M.I.T., 1985.

[7] R. Gallager, "A perspective on multiaccess channels," *IEEE Transactions on Information Theory*, vol. 31, pp. 124–142, Mar. 1985.

[8] M. S. Pinsker, "On the complexity of decoding," *Problemy Peredachi Informatsii*, vol. 1, no. 1, pp. 84–86, 1965.

[9] J. Massey, "Capacity, cutoff rate, and coding for a direct-detection optical channel," *IEEE Transactions on Communications*, vol. 29, pp. 1615–1621, Nov. 1981.

[10] E. Arıkan, "On the origin of polar coding," *IEEE Journal on Selected Areas in Communications*, vol. 34, pp. 209–223, Feb. 2016.

[11] P. Elias, "Error-free coding," *Transactions of the IRE Professional Group on Information Theory*, vol. 4, pp. 29–37, Sept. 1954.

[12] H. Imai and S. Hirakawa, "A new multilevel coding method using error-correcting codes," *IEEE Transactions on Information Theory*, vol. 23, pp. 371–377, May 1977.

[13] E. Arıkan, "Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels," *IEEE Transactions on Information Theory*, vol. 55, pp. 3051–3073, July 2009.

[14] R. Mori and T. Tanaka, "Performance of polar codes with the construction using density evolution," *IEEE Communications Letters*, vol. 13, pp. 519–521, July 2009.

[15] R. Pedarsani, S. H. Hassani, I. Tal, and E. Telatar, "On the construction of polar codes," in *2011 IEEE International Symposium on Information Theory Proceedings*, pp. 11–15, IEEE, July 2011.

[16] I. Tal and A. Vardy, "How to construct polar codes," *IEEE Transactions on Information Theory*, vol. 59, pp. 6562–6582, Oct. 2013.

[17] E. Arıkan and E. Telatar, "On the rate of channel polarization," in *2009 IEEE International Symposium on Information Theory Proceedings*, pp. 1493–1495, IEEE, June 2009.

[18] V. Guruswami and P. Xia, "Polar codes: speed of polarization and polynomial gap to capacity," *IEEE Transactions on Information Theory*, vol. 61, pp. 3–16, Jan. 2015.

[19] Y. Polyanskiy, H. Poor, and S. Verdú, "Channel coding rate in the finite blocklength regime," *IEEE Transactions on Information Theory*, vol. 56, pp. 2307–2359, May 2010.

[20] I. Tal and A. Vardy, "List decoding of polar codes," in *2011 IEEE International Symposium on Information Theory Proceedings*, pp. 1–5, July 2011.

[21] R. Fano, "A heuristic discussion of probabilistic decoding," *IEEE Transactions on Information Theory*, vol. 9, pp. 64–74, Apr. 1963.

[22] I. Reed, "A class of multiple-error-correcting codes and the decoding scheme," *Transactions of the IRE Professional Group on Information Theory*, vol. 4, pp. 38–49, Sept. 1954.

[23] D. E. Muller, "Application of Boolean algebra to switching circuit design and to error detection," *Transactions of the I.R.E. Professional Group on Electronic Computers*, vol. EC-3, pp. 6–12, Sept. 1954.

[24] N. Seshadri and C. E. W. Sundberg, "List Viterbi decoding algorithms with applications," *IEEE Transactions on Communications*, vol. 42, pp. 313–323, Feb. 1994.

[25] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*. Upper Saddle River, NJ, USA: Prentice Hall Press, 3rd ed., 2009.

Tooting Our Horns: Practicing and Preparing Research Pitches

Gireeja Ranade and Christina Lee Yu

The first event we hosted as co-chairs of the Women in Information Theory Society (WITHITS) was a short lunch workshop at ISIT 2019 which focused on developing research pitches. Communicating our work and sharing our ideas with colleagues is a very important part of being a researcher, and yet this is a skill that is not often formally taught. In the process of brainstorming for this event, we reflected upon our journeys from graduate students to faculty, and recalled feeling nervous when describing and presenting one's research identity in casual conference networking events. Our aim with this workshop was to give participants an opportunity to practice communicating and promoting their own work in a safe and encouraging environment. As an additional goal, we hoped that being able to connect and meet other women and mentors on the first day of the conference would catalyze further conversations and provide relationships of support that would build confidence for participants to fully enjoy and engage in productive research conversations through the remainder of the conference.

The workshop started off with Tara Javidi sharing some guidelines on what makes a good research pitch. Plenary speaker Muriel Medard as well as Daniela Tuninetti and Lalitha Sankar shared examples of pitches as well—giving examples in different settings, such as pitches to a funding agency, pitches about a conference talk, or pitches to a friend or family member. After hearing advice and examples, the rest of the event was dedicated to practicing in small groups to answer the following prompt: Suppose that you meet a new person at coffee break that greeted you with "Hello, nice to meet you! What do you work on?" The participants took turns to practice their response and give each other constructive feedback on their pitches. Senior members of the community were present to give feedback as well.

There were more than 120 people who attended the event. We received very positive feedback for the event—some groups discovered new research connections in their work, some participants were inspired to replicate the same event at their own university or lab group meetings, and other participants even commented that we should hold the exact same event every single year! Going forward, we are excited for hosting future events that support the mission of WITHITS—addressing the needs of and encouraging the participation of our underrepresented demographics, while being of interest and use to the community at large. We envision events that provide mentoring and build a support network, and provide training and practice for other similarly important but under-taught skills for flourishing as an academic.

11th Asia-Europe Workshop (AEW11) on "Concepts in Information Theory and Communications"

Organizers: Han A.J. Vinck and Kees A. Schouhamer Immink

The 11th Asia-Europe workshop on "Concepts in Information Theory and Communications" (AEW11) was held in Rotterdam-Netherlands on July 3-5, 2019. Thirty-three participants enjoyed the beautiful venue of the clubhouse of the Royal Maas Rowing and Sailing Club at the quay of a branch of the river Rhine, called the Maas in the very center of Rotterdam.

The workshop is based on a long-standing cooperation between Asian and European scientists. The very first workshop was held thirty years ago in Eindhoven, the Netherlands in 1989. The main idea of the workshop is threefold: 1) improvement of the communication between scientists in different parts of the world; 2) exchange

of knowledge and ideas; 3) pay a tribute to a well-respected and special scientist.

For this workshop, Hiroyoshi Morita accepted the invitation to be the guest of honor and to be the key lecturer. Hiro is a well-known information theorist with many original contributions. We have also appreciated very much his many contributions to the Information theory community in general. Hiro gave an overview on 'Antidictionary and its Applications'.

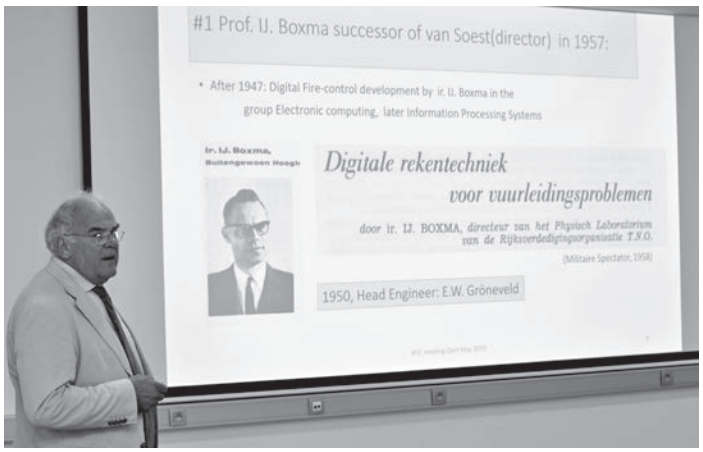
The other sixteen presentations showed examples of concepts of error correcting codes, time series analysis, cryptography, multi-server load balancing, convolutional codes, and coding for memories. The proceedings of the workshop can be found at https://arxiv.org/abs/1907.02944.

The workshop started with an informal get-together on Wednesday evening. On Thursday evening, after a reception in the garden near the river, a delicious dinner was served for attendees and spouses in the clubhouse's great hall. The workshop was concluded with a boating and hiking trip to Kinderdijk, a well-preserved Dutch polder with nineteen original windmills that have kept the Dutch feet dry for many centuries.

40th Symposium on Information Theory in the Benelux

Jos Weber

The 2019 Symposium on Information Theory and Signal Processing in the Benelux took place at the KU Leuven Technology Campus in Ghent, Belgium, on May 28 and 29. It was extremely well organized by Liesbet Van der Perre, Sofie Pollin, Alexander Bertrand, Gilles Callebaut, Bert Cox, and Kevin Verniers, all from the Katholieke Universiteit Leuven. The long range of annual symposia coordinated under the auspices of the "Werkgemeenschap voor Informatie- en Communicatietheorie (WIC)" started in 1980. Later also the IEEE Benelux Information Theory Chapter became involved, and since 2011 the symposia are co-organized with the IEEE Benelux Signal Processing Chapter. The goal of the conference is to bring together researchers from academia and industry within the Benelux countries (Belgium, Netherlands, and Luxemburg), to share ideas, problems and solutions relating to the multifaceted aspects of signal processing and information theory.



The organizing committee welcomed 76 attendees to the 40th edition of this symposium. In order to celebrate the jubilee, several special events were included into the program. There were keynote lectures by (former) WIC chairmen: Han Vinck (University of Duisburg-Essen, WIC chair 1998–2001) talked about "Information Theory and Memory Systems", Peter de With (TU Eindhoven, WIC chair 2001–2006) presented "Flying through WIC Benelux history: from video coding towards image analysis", while Jos Weber (TU Delft, WIC chair 2006-present) discussed "Channel Coding in the Benelux". Furthermore, a special session on "Searching life-critical information, early computer-aided detection of cancer" was organized. To enrich the social cohesion, the participants were guided through the historic city center of Ghent, entertained with a big band concert, a food truck, and surprised with birthday cakes at the conference dinner.

As usual, the core of the symposium was formed by oral and poster presentations, mainly by PhD and MSc students. Different

research fields were addressed, from quantum cryptography to biomedical signal processing to signal detection and estimation. The best student paper award was won by Simon Geirnaert (KU Leuven) for "Expected Switching Time: a Markov Chain Based Performance Metric to Evaluate Auditory Attention Decoding Algorithms", joint work with T. Francart and A. Bertrand. The best student presentation award was won by Miao Sun (TU Delft), for "Atrial Activity Extraction Based on Graph-Time Signal Processing", joint work with E. Isufi, N.M.S. de Groot, and R.C. Hendriks.

At the WIC General Assembly that took place during the symposium, Ludo Tolhuizen from Philips Research was awarded the honorary membership of the WIC for his extensive service to the WIC. In particular, he has served as a board member since 1998. During the years 1998-2006 and 2010-2019 he was the WIC secretary.

The symposium proceedings and slides of the keynotes are available via w-i-c.org. At this web site also future events will be announced. The 2020 Symposium on Information Theory and Signal Processing in the Benelux will be organized by TU Eindhoven, dates to be decided. Information theory in the Benelux is alive and kicking!

The Fifth London Symposium on Information Theory (LSIT)

Osvaldo Simeone and Deniz Gündüz

Two weeks after the end of post-war soap rationing, and a month after BBC's first overseas live TV broadcast (from France), a distinguished group of academics gathered at the Royal Society in London to talk about information theory. It was 1950—only two years after the publication of Shannon's seminal paper and of Wiener's "Cybernetics"—and the new ideas of information, control, and feedback were quickly making their way from engineering to the natural, social, and human sciences, begetting new insights and raising new questions.

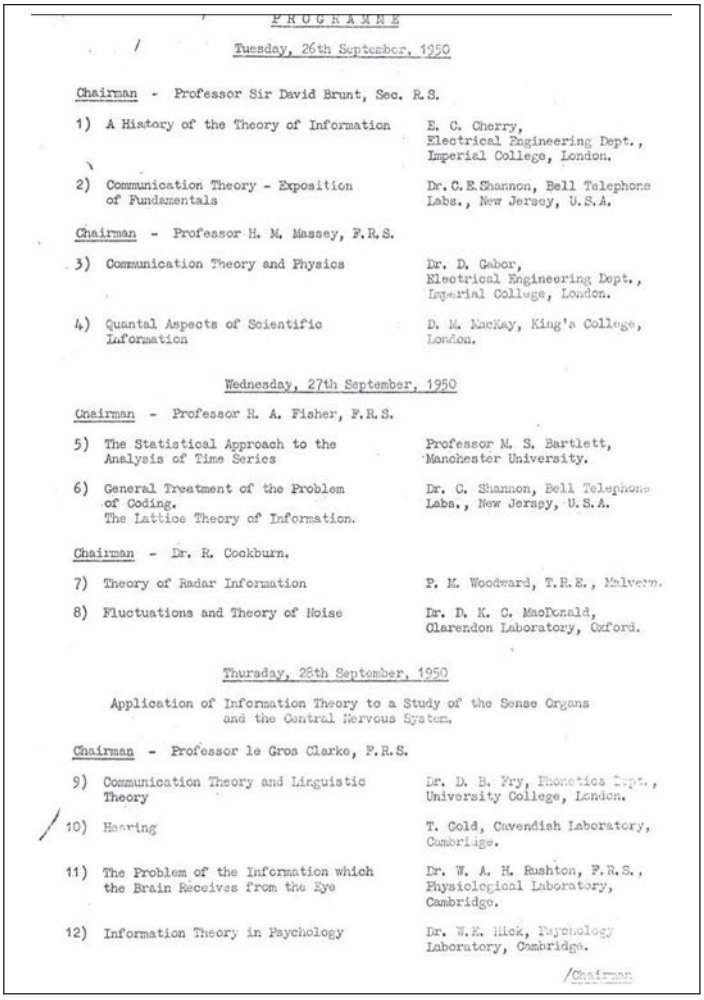
This "cybernetic moment" [2] underpinned the first four editions of the London Symposium on Information Theory (LSIT), with the first meeting in 1950 followed by the symposia in 1952, 1955, and 1960. The program in 1950, shown in Fig. 1, featured two talks by Shannon on communications and coding, as well as a number of presentations on topics ranging from physics, statistics, and radar, to linguistics, neuroscience, psychology, and neurophysiology. The first LSIT was also notable for two written contributions by Alan Turing, who could not attend in person. One of the contributions offered the following ante-litteram definition of machine learning:

"*If [...] the operations of the machine itself could alter its instructions, there is the possibility that a learning process could by this means completely alter the programme in the machine.*"

According to the report [2], the second meeting, held in 1952, was characterized by an "*emphasis on the transmission and analysis of speech*", while the third LSIT in 1955 covered again a large variety of topics, including "*anatomy, animal welfare, anthropology, [...] neuropsychiatry, [...] phonetics, political theory*". David Slepian, one of the participants, would later write in his Bell Labs report about this third meeting that the "*best definition I was able to get as to what constituted 'The Information Theory' was 'the sort of material on this program'*!" [2]. At the same time, the heterogeneity of topics in the program may have been one of the motivations behind Shannon's "Bandwagon" paper published the following year [3]. In it, Shannon famously warned against the indiscriminate application of information theory based solely on the abstract relevance of the concept of information to many scientific and philosophical fields.

The fourth LSIT was held in 1960 and featured among its speakers Marvin Minsky, one of the founding fathers of Artificial Intelligence (AI), who delivered a talk entitled "Learning in Random Nets".

In the middle of our own "AI moment", the time seemed right to bring back to London the discussion initiated in the fifties and sixties during the first four LSIT editions. And so, with a temporal leap of almost sixty years, the fifth LSIT was held at King's College London on May 30-31, 2019. The symposium was organized



PROGRAMME

Tuesday, 26th September, 1950

Chairman - Professor Sir David Brunt, Sec. R.S.

1) A History of the Theory of Information — E. C. Cherry, Electrical Engineering Dept., Imperial College, London.

2) Communication Theory - Exposition of Fundamentals — Dr. C. E. Shannon, Bell Telephone Labs., New Jersey, U.S.A.

Chairman - Professor H. M. Massey, F.R.S.

3) Communication Theory and Physics — Dr. D. Gabor, Electrical Engineering Dept., Imperial College, London.

4) Quantal Aspects of Scientific Information — D. M. MacKay, King's College, London.

Wednesday, 27th September, 1950

Chairman - Professor R. A. Fisher, F.R.S.

5) The Statistical Approach to the Analysis of Time Series — Professor M. S. Bartlett, Manchester University.

6) General Treatment of the Problem of Coding. The Lattice Theory of Information. — Dr. C. Shannon, Bell Telephone Labs., New Jersey, U.S.A.

Chairman - Dr. R. Cockburn.

7) Theory of Radar Information — P. M. Woodward, T.R.E., Malvern.

8) Fluctuations and Theory of Noise — Dr. D. K. C. MacDonald, Clarendon Laboratory, Oxford.

Thursday, 28th September, 1950

Application of Information Theory to a Study of the Sense Organs and the Central Nervous System.

Chairman - Professor le Gros Clarke, F.R.S.

9) Communication Theory and Linguistic Theory — Dr. D. B. Fry, Phonetics Dept., University College, London.

10) Hearing — T. Gold, Cavendish Laboratory, Cambridge.

11) The Problem of the Information which the Brain Receives from the Eye — Dr. W. A. H. Rushton, F.R.S., Physiological Laboratory, Cambridge.

12) Information Theory in Psychology — Dr. W. E. Hick, Psychology Laboratory, Cambridge.

/Chairman

Figure 1. Program of the first LSIT, held in 1950.

by Deniz Gündüz and Osvaldo Simeone from Imperial College London and King's College London, respectively—two institutions that featured prominently in the first editions of LSIT (see Fig. 1).

While heeding Shannon's warning, the program of the symposium aimed at exploring the "daisy" of intersections of machine learning with fields such as statistics, machine learning, physics, communication theory, and computer science. Each day featured two keynote talks, along with two invited session, and a poster session with invited as well as contributed posters submitted to an open call. The first day was devoted to the intersection between machine learning and information theory, while the second day focused on novel applications of information theory.

The first day started with a keynote by Michael Gastpar (EPFL), who presented a talk entitled "Information measures, learning and generalization". This was followed by an invited session on "Information theory and data-driven methods", chaired by Iñaki Esnaola (University of Sheffield), which featured talks by Bernhard Geiger (Graz University of Technology) on "How (not) to train your neural network using the information bottleneck principle"; by Jonathan Scarlett (National University of Singapore) on "Converse bounds for Gaussian process bandit optimization"; by Changho Suh (KAIST) on "Matrix completion with graph side information"; and by Camilla Hollanti (Aalto University) on "In

Figure 2. General co-chairs with some of the student and postdoc volunteers.

the quest for the capacity of private information retrieval from coded and colluding servers". The session was interrupted by a fire alarm that was carefully timed by the organizers in order to give the attendees more time to enjoy the storied surrounding of the Strand Campus of King's College London. After lunch, the symposium kicked off with a keynote talk by Phil Schniter (Ohio State University) on "Recent advances in approximate message passing", which was followed by an invited session on "Statistical signal processing", organized by Ramji Venkataramanan (University of Cambridge), which featured talks by Po-Ling Loh (University of Wisconsin-Madison) on "Teaching and learning in uncertainty"; by Cynthia Rush (Columbia University) on "SLOPE is better than LASSO"; by Jean Barbier (EPFL) on "Mutual information for the dense stochastic block model: A direct proof"; and by Galen Reeves (Duke University) on "The geometry of community detection via the MMSE matrix". The first day was ended by a poster session organized by Bruno Clerckx (Imperial College London); by wine, refreshments, and by the view on the Thames and the Waterloo bridge from the 8th floor of the Bush House.

The second, and last day, started off with a keynote by Kannan Ramchandran (Berkeley) on "Beyond communications: Codes offer a CLEAR advantage (Computing, LEArning, and Recovery)". Next was an invited session on "Information theory and frontiers in communications", chaired by Zoran Cvetkovic (King's College London), with talks by Ayfer Özgür (Stanford) on "Distributed learning under communication constraints"; by Mark Wilde (Louisiana State University) on "A tale of quantum data processing and recovery"; by Michèle Wigger (Telecom ParisTech) on "Networks with mixed delay constraints"; by Aaron Wagner (Cornell University) on "What hockey and foraging animals can teach us about feedback communication"; and by Ofer Shayevitz (Tel Aviv University) on "The minimax quadratic risk of distributed correlation estimation". The afternoon session was opened by Yiannis Kontoyiannis (University of Cambridge), who gave a keynote on "Bayesian inference for discrete time series using context trees", and continued with an invited session on "Post-quantum cryptography", organised by Cong Ling (Imperial College London), with talks by Shun Watanabe (Tokyo University of Agriculture and Technology) on "Change of measure argument for strong converse and application to parallel repetition"; Qian

Guo (University of Bergen) on "Decryption failure attacks on post-quantum cryptographic primitives with error-correcting codes"; Leo Ducas (CWI) on "Polynomial time bounded distance decoding near Minkowski's bound in discrete logarithm lattices"; and Thomas Prest (PQShield) on "Unifying leakage models on a Rényi day". As the first, the second was not a rainy day and attendees were able to enjoy the view from the Bush House terrace with wine and mezes, while discussing results from poster sessions organised by Mario Berta (Imperial College London), and Kai-Kit Wong (University College London).

Videos of all talks are availabe on YouTube (https://tinyurl.com/y5w92rga).

Registration was free and more than 150 students, researchers, and academics were in attendance. Support was provided by the European Research Council (ERC) under the European Union's Horizon 2020 Research and Innovation Programme (Grant Agreements No. 725731 and 677854).

LSIT has outlived the cybernetic movement, and it may well continue beyond the current "AI moment". The organizers hope that we will not wait for another 60 years for the next information theory meeting in London, and would like LSIT to become a regular biennial meeting of the information theorists and researchers from related fields in London, in alternating years with the International Zurich Seminar, which has so far been the only regular meeting of its kind in Europe. Coincidentally, the first ever international meeting on information theory was held in Zurich in September 11–22, 1950, only one week before the London symposium, during the congress of the International Union of Radio Science (URSI) [4].

References

[1] N. Blachman, "Report on the third London Symposium on Information Theory," *IRE Transactions on Information Theory*, vol. 2, no. 1, pp. 17–23, March 1956.

[2] R. R. Kline, "The Cybernetics Moment Or Why We Call Our Age the Information Age", John Hopkins University Press, 2015.

[3] C. E. Shannon, "The bandwagon", *IRE Transactions on Information Theory*, vol. 2, no. 1, Mar. 1956.

[4] F. L. H. M. Stumpers, "Informatietheorie: Een terugblik op de eerste periode" Tijdschrift van het Nederlands Elektronica- en Radiogenootschap deel 58, nr. 2, pp. 65–71, 1993.

In Memoriam: Robert J. McEliece (1942–2019)

Dariush Divsalar and Mario Blaum

Tribute to Robert J. McEliece, who passed away May 8, 2019. Dariush Divsalar and Mario Blaum write about Bob's broad and substantial contributions to information theory, coding theory and cryptography.

Robert J. McEliece was born in Washington, DC, on May 21, 1942, and passed away on May 8, 2019 in Pasadena. He received the B.S. and Ph.D. degrees in mathematics from the California Institute of Technology in 1964 and 1967, respectively, and attended Trinity College, Cambridge University, England, during 1964–65. From 1963 to 1978 he worked at Caltech's Jet Propulsion Laboratory (JPL), and he had been a consultant at JPL since 1978. From 1978 to 1982 he was Professor of Mathematics and Research Professor at the Coordinated Science Laboratory, University of Illinois, Urbana-Champaign. He joined the faculty at Caltech in 1982, and became an Allen E. Puckett Professor in 1997. He was also a regular consultant at the Sony Corp. in Tokyo.

Prof. McEliece made fundamental contributions to the theory, design and practice of channel codes for communication systems. Prof. McEliece's achievements are many. At Jet Propulsion Laboratory, Prof. McEliece has contributed to the design and analysis of many coded interplanetary telecommunication systems, for example the Golay-coded non-imaging system for the Voyager spacecraft, and the "Big Viterbi Decoder" which has been used on the Galileo, Mars Pathfinder, Cassini, and Mars Exploration Rover missions. He has won several NASA awards for this work.

As a faculty member at Caltech, he has five times won awards for excellence in teaching, and has mentored more than 30 Ph.D. students, four of whom are IEEE Fellows. From 1990–1999, he served as Executive Officer (chairman) for the Electrical Engineering Department, and under his leadership Caltech's small (12 FTE) EE Department rose to rank 5th nationally, behind only MIT, Stanford, Berkeley, and the University of Illinois.

Prof. McEliece is the author of three textbooks and more than 250 research articles, jointly with more than 75 coauthors (his top 12 papers yield 10291 citations on Google Scholar— Prof. McEliece is currently listed as a Highly Cited Researcher by Thompson-ISI).

Besides his technical achievements, Prof. McEliece had many other interests. Among them, he was an avid runner, and he loved music and singing. He had an affable personality and he was loved and respected by his peers and his students. He influenced greatly the careers of many of us. He was a truly outstanding lecturer and was an excellent popularizer of information theory. He can be described as a "mathemagician". Enjoy for example his lecture Safety in Numbers–Protecting Data Mathmagically, Part 1 and Part 2 on YouTube.

He is survived by three daughters, a son and a step-daughter.

Below is a chronological list of some of Prof. McEliece's most important achievements.

1) McEliece's Theorem. This theorem, which identifies the largest power of p that divides all the weights in a p-ary cyclic code, and which contains the celebrated Ax divisibility theorem as a special case, is one of the deepest mathematical results to come out of coding theory. McEliece's theorem has inspired a large and impressive body of later work by Wilson, Calderbank, Katz, and others. Reference [R1] and [R2].

2) The Theory of Information and Coding. In print continuously since 1977, this classic textbook book has been compared to Richard Feynman's Lectures on Physics, as a standard and authoritative book in its field. Reference [R3]

3) The JPL Bound. Since 1977 this result has stood as the best known upper bound on the basic combinatorial problem of Information Theory: the tradeoff between rate and minimum distance of the best binary codes. Winner of an Information Theory Society Golden Jubilee Award, 1998. Reference [R4]

4) The McEliece Public-key Cryptosystem. Has withstood repeated continuous attacks of cryptanalysts for more than 30 years, and thus (with RSA) is one of a small handful of successful public-key cryptosystems. McEliece cryptosystem is a candidate for post-quantum cryptography since it is immune to attacks using Shor's algorithm. Reference [R5]

5) Decoding is NP-hard. The first proof that maximum-likelihood decoding of linear block codes is an intractable problem. This result has inspired many similar results by later researchers. Reference [R6]

6) Block interference channel models. A class of channel models with memory that is (a) simple enough to allow precise analysis and (b) realistic enough to yield insights into real channels. This class of channel models has proved essential in the study of wireless fading channels. Reference [R7]

7) The capacity of the Hopfield Neural network. This paper gave the first rigorous estimate of the potential of neural-network type memories. Reference [R8]

8) Turbo decoding and belief propagation. Winner of the 1998 Leonard G. Abraham award. This paper put the term "belief propagation" in the coding theory vocabulary. Reference [R9]

9) The Generalized Distributive Law. An important synthesis showing deep and previous unsuspected connections between the fast Fourier Transform, Viterbi's algorithm, Turbo decoding, and many other basic algorithms. (a patent). Reference [R10]

10) Repeat-Accumulate Codes. An astonishingly simple and powerful class of codes which bridge the gap between turbo-codes and LDPC codes, and which have become an industry standard. (two patents). Reference [R11] and [R12].

References

[R1] Weight Congruences for p-ary Cyclic Codes, Discrete Math. 3 (1972), pp.177–192.

[R2] Zeroes of Functions in Finite Abelian Group Algebras (with P. Delsarte), Am. J. Math. 98 (1976), pp. 197–224.

[R3] The Theory of Information and Coding (Addison-Wesley, 1977); 2nd Ed., (Cambridge U Press, 2002); Student Ed., (Cambridge U Press, 2004).

[R4] New Upper Bounds on the Rate of a Code via the Delsarte-MacWilliams Inequalities (with E. Rodemich, H. Rumsey, L. Welch), IT-23 (1977), pp. 57–166.

[R5] A Public-Key Cryptosystem Based on Algebraic Coding Theory, JPL DSN Progress Reports vol. 44 (1978), pp. 114–116.

[R6] On the Inherent Intractability of Certain Coding Problems (with E.R. Berlekamp and H. Van Tilborg), IT-24 (1978), pp. 384–386.

[R7] Channels with Block Interference (with W. Stark), IT-30 (1984), pp. 44–53.

[R8] The Capacity of the Hopfield Associative Memory (with E. C. Posner, E. Rodemich, and S. Venkatesh), vol. IT-33 (July 1987), pp. 461–482. Reprinted in: V. Vemuri, Ed., Artificial Neural Networks: Theoretical Concepts. Los Angeles, IEEE Computer Society Press, 1988.

[R9] Turbo Decoding as an Instance of Pearl's 'Belief Propagation' Algorithm,' (with David MacKay and J.-F. Cheng), IEEE J. Sel. Areas Comm., vol. 16, no. 2 (Feb. 1998), pp. 140–152.

[R10] The Generalized Distributive Law (with S. M. Aji), IEEE Trans. Inform. Theory, vol. IT-46, no. 2 (March 2000), pp. 325–343.

[R11] Coding Theorems for 'Turbo-Like' Codes, (with D. Divsalar and H. Jin), Proc. 1998 Allerton Conference, pp. 201–210.

[R12] "Irregular Repeat-Accumulate Codes," (with H. Jin and A. Khandekar), Proc. 2nd. International Conf. Turbo Codes, Brest, France, 4–7 Sept. 2000, pp. 1–8.

References note: "IT" is short for the IEEE Transactions on Information Theory.

Robert J. McEliece's Honors/Awards

1) Associated Students of the California Institute of Technology Award for Excellence in Teaching, 1985, 1989, 1990, 1999.

2) Caltech Graduate Student Council Teaching Award, 1996.

3) Erdos number one: "Ramsey Bounds for Graph Products" (with Paul Erdos and Herbert Taylor), Pac. J. Math. 37 (1971), pp. 45–46.

4) NASA Group Achievement Award for Voyager Mission Operations Systems Design and Development, June 1981.

5) Elected President of the IEEE Information Theory Group (one-year term of office, 1984)

6) NASA Group Achievement Award for the Advanced Error-Correcting Code Research and Development Team, June 1992. ("In recognition of research and development resulting in a new error-correcting system providing an increase of data rate by a factor of 1.6 to benefit all future space missions")

7) Elected to National Academy of Engineering, 1998.

8) Elected Fellow, IEEE, 1984

9) Paper "New Upper Bounds on the Rate of a Code via the Delsarte-MacWilliams Inequalities," selected for an Information Theory Society Golden Jubilee Award, 1998

10) Paper "Turbo Decoding as an Instance of Pearl's 'Belief Propagation' Algorithm" awarded the 1998 Leonard G. Abraham Prize paper Award

11) IEEE Third Millennium Medal.

12) IEEE Information Theory Society 2004 Claude E Shannon Award

13) IEEE Alexander Graham Bell Medal, 2009

Read more about R. J. McEliece's life at Caltech News, https://news.berkeley.edu/2019/04/18/elwyn-berlekamp-game-theorist-and-coding-pioneer-dies-at-78/

In Memoriam: Elwyn Berlekamp (1940–2019)

Jim Omura

Elwyn Berlekamp passed away on April 9, 2019 in Berkeley, California. He was a brilliant mathematician and engineer who did groundbreaking research in Information Theory and Combinatorial Game Theory. He was also a successful entrepreneur, helped create a highly profitable quantitatively managed fund, served in leadership roles for academic societies, and was a generous supporter of STEM education and mathematics research. Elwyn is survived by his wife, Jennifer; daughters Persis, an art historian at the University of Chicago, and Bronwen Berlekamp O'Wril of Portland, Maine; and son David of Oakland.

Early Years

Elwyn Berlekamp was born on September 6, 1940 in Dover, Ohio. His family moved to Northern Kentucky, where Elwyn was class president at Highland High School. In 1958 he and I entered MIT as freshmen. I first met Elwyn in our sophomore year when we played together on the East Campus intramural football team.

Elwyn was the smartest student I met at MIT. He took extra courses during his undergraduate years, completing his BS and MS degrees in 4 years and his PhD two years later, all at MIT. In addition, he taught himself Russian and together with other undergraduate students developed the first chess playing software program, which is featured in the Computer History Museum in Mountain View, California. Although an Electrical Engineering student, he was among five winners of the national Putnam Mathematics competition in December 1961. His PhD thesis, *Block Coding with Noiseless Feedback*, spawned several publications. One key idea was to recast the problem as an asymmetric combinatorial game between the Coder and the Noisemaker and then to find asymptotically optimum strategies for playing that game. This was followed by his groundbreaking research work in Information Theory, Mathematics, and Combinatorial Game Theory. Less known is his mostly classified consulting work on cryptographic research for the Institute for Defense Analysis (IDA) in Princeton, New Jersey.

Research in Information Theory

In addition to his long career at the University of California Berkeley, Elwyn worked for short periods at Bell Labs, the Jet Propulsion Laboratory (JPL), and IDA.

He invented a series of algorithms and related implementations which made powerful error-correcting codes useful in many applications. Best-known among these results was his algorithm to factor polynomials over finite fields and the Berlekamp-Massey algorithm to find linear recursions in long data streams. These were published in his 1968 book, *Algebraic Coding Theory*. Another major new result in that book was the enumeration of the number of information bits in long binary BCH codes. This book won the IEEE Information Theory Group's annual best research paper award. The applicability of some of these algorithms to problems in cryptography attracted the attention of the National Security

Agency (NSA), who in 1967-1968 recruited Elwyn to become a consultant to their research group at IDA.

Entrepreneur
Some believed that Elwyn's algorithms were impractical. Without angel investors or venture capital, he founded Cyclotomics to develop commercial implementations, starting with the world's foremost Galois Field computer which shattered the then-prevalent myth that powerful high-speed algebraic error-correction was not feasible. Elwyn's Galois Field computer is now in the Computer History Museum.

When I met him during this period, I was amazed to see Elwyn programming an early microprocessor and designing integrated circuits. This work led to Elwyn becoming the youngest member of the National Academy of Engineering in 1977.

Cyclotomics bootstrapped its growth with profitable sales, and eventually grew to a peak of 40 people. It worked with NASA to design and build the error-correction decoders for the downlink of the Hubble Space Telescope. Beyond communications, Cyclotomics also pioneered applications of algebraic error-correction technology to several data storage technologies, including magnetic tapes, magnetic discs, and optical discs. By the mid-1980s, there were over 40 US companies trying to develop read-write optical memories. Cyclotomics developed controllers for several of them. The biggest of these companies was Eastman Kodak which acquired Cyclotomics in 1985.

In 1984 Cyclotomics spun off its consulting contracts in cryptography to a new startup, which became known as Cylink and which I co-founded with Elwyn and others. Cylink obtained angel funding from a group formed by Jim Simons, whom Elwyn had met at an IDA interview in 1967. In 1996 Cylink went public on the NASDAQ stock exchange.

The Medallion Quant Fund
Simons had started a hedge fund management firm, which became Renaissance Technologies in 1982. Six years later Renaissance established the Medallion quant fund using Leonard Baum's mathematical models, which were improved by pioneering algebraist James Ax to explore correlations from which they could profit. Around 1989 this fund, then called Axcom with Ax as CEO, was not doing well. Having met again as members of Cylink's board of directors, Simons turned to Elwyn to run Medallion from Berkeley. Elwyn bought out most of Ax's stake in Axcom and became its CEO. Over a six-month period, he worked with Simons, Sandor Straus, and consultant Henry Laufer to overhaul Medallion's trading system. In 1990 Elwyn led Medallion to a 55.9% gain, net of fees, and then returned to teaching at Berkeley after selling out his Axcom shares to Simons at six times the price he had paid 16 months earlier. Straus took the reins of Medallion's revamped trading system, and Medallion returned 39.4% in 1991, 34% in 1992 and 39.1% in 1993. They continued hiring mathematicians, engineers, and scientists and expanded into trading stocks as well as futures. The Medallion fund became the most successful hedge fund ever, averaging a 71.8% annual return, before fees, from 1994 through mid-2014. It made Simons the leading fund manager on Wall Street.

In the early 2000s, four new partners and Elwyn launched another quantitatively managed fund, called Berkeley Quantitative (BQ). After two years of development and studies, it accepted money from outside investors and began trading. As more investors entered, the fund size grew. Net performance reached 17%, and the fund size reached $250 million about 1.5 years after trading had begun. But the next couple of months saw unfriendly market conditions aggravated by some internal organizational problems and BQ was closed down after paying off its debts. The initial investors realized a net return of about 2% after 3 years. Elwyn felt that BQ was not a success, but it did much better than the wipe-outs experienced by many startups at that time.

Combinatorial Game Theory
Elwyn's fascination with mathematical games began when he learned to play *Dots-and-Boxes* in the first grade. Years later he discovered several mathematical theorems that underlie this game and others. In mathematics, his best-known work was in Combinatorial Game Theory, partly disseminated in his four-volume work *Winning Ways* with John H. Conway and Richard Guy. According to Martin Gardner, author of the extremely popular *Mathematical Games* column in *Scientific American* from 1957 to 1982, *Winning Ways* was the "greatest contribution of the 20th century to the burgeoning field of recreational mathematics. No other work has been so packed with completely new and significant material, or presented with so much wit, depth, and clarity." These books also became the foundation of a less-recreational subject called *Combinatorial Game Theory*. It was officially accepted by *Math Reviews* (now *MathSciNet*) as a new branch of mathematics and attracted the interest and contributions of many mathematicians and computer scientists.

Another of Elwyn's accomplishments in Combinatorial Game Theory was his analysis of positions in the endgame of Go. With David Wolfe, he published the book *Mathematical Go*. He demonstrated the effectiveness of his theory by setting up a plausible endgame position from which he beat one of the Japanese champions of the game, after which he set up the same position, reversed the board, and beat the master a second time. He also invented a variation of the game called *Coupon Go*, which is closer to elegant mathematical theories. This attracted the attention of both mathematicians and several world-class professional Go players.

Contributions to STEM education and the Mathematical Sciences Research Institute (MSRI)
Elwyn's father was a minister, and one sees the father's influence on the son in a strong and consistent ideal of service to the greater good.

Elwyn was active in the popularization of Science, Technology, Engineering, and Mathematics (STEM), directed both at K-12 education and at adults. He served on the governing boards of each of the two foremost private schools in Oakland, California, College Preparatory School and Head-Royce. He also directed his efforts towards extra-curricular education. In the late 1970s, he helped finance the start of the Berkeley Math Circle for junior high school students who met in a classroom one evening per week to share and enjoy solving math or logic problems not covered in their school curricula. He felt that one key to the success of such efforts was to eradicate (or at least blur) the line many students imagine between "real" mathematics/engineering and "recreational mathematics."

In 1979 Elwyn joined other MSRI founders in a meeting that persuaded UC Berkeley Chancellor Albert Bowker to support (verbally but not financially) MSRI as an allied but independent

off-campus entity, with its own board of governors, neither controlled nor overseen by UC Berkeley. Serving as Chairman of the Board from 1994 to 1997, Elwyn hired the current MSRI director, David Eisenbud who wrote of Elwyn's connections to MSRI in his *Memoriam: Elwyn Berlekamp*.

Starting in 2015 he devoted much of his time to the preparation of short introductory videos aimed at getting more junior high school students interested in combinatorial games and some of the mathematics behind them. In 2017 he also became the initial donor to fund "America Asks, Science Answers", a new public information campaign by the National Academies.

With his wife Jennifer, Elwyn supported various charitable causes and in 2013 founded the Elwyn and Jennifer Berlekamp Foundation, a small private foundation based in Oakland to support math and science outreach and education in general and Combinatorial Game Theory in particular.

In Closing

Like others with Top Secret clearances in cryptography, Elwyn refrained from publishing any papers directly relating to cryptography, whether they were classified or not. We once visited the National Security Agency (NSA) where I was able to see classified reports he had authored. In the 1970s academic papers on cryptography started to appear in academic journals, arousing concerns at the NSA that this information might be useful to U.S. adversaries. To address these concerns, the Director of NSA invited Elwyn to meetings with leading academics. Elwyn later convened similar meetings at UC Berkeley which led to a commission of relevant scholarly societies addressing these concerns.

Throughout his career Elwyn continuously made significant contributions to research and teaching, developed practical applications as a successful entrepreneur, supported education in mathematics, and held leadership roles in academic societies. He served on over 45 boards of various kinds. In 1973, he became one of a dozen faculty co-founders of the Computer Science Division within UC Berkeley's Department of Electrical Engineering and Computer Science. His many students have in turn made significant impacts in electrical engineering and computer science. For example, one of his students, Ken Thompson, became the co-inventor of the Unix operating system. Many others have become leaders in academia and industry. Elwyn served in numerous leadership roles at the University of California, IEEE, the National Academy of Sciences, the National Academy of Engineering, non-profit educational organizations, and MSRI.

Elwyn and I first met at MIT sixty years ago. Over the past few years we would meet for long lunches once every couple of months, sharing stories and often discussing world events and social issues. Elwyn cared a lot about education and worried about inequalities in our society. It has been a privilege to have had Elwyn as a colleague and friend.

Additional Information about the Life and Legacy of Elwyn Berlekamp

- Elwyn Berlekamp website: https://www.ejbf.org/elwyn-home (This website also contains his lectures including at UCB in 2006, Kailath Lecture at Stanford, and Viterbi lecture at USC.)
- Computer History Museum: https://www.computerhistory.org/chess/orl-433444ecc827d/
- MSRI: http://www.msri.org/web/msri/communications/elwyn-berlekamp
- Wikipedia page: https://en.wikipedia.org/wiki/Elwyn_Berlekamp
- Berkeley News: Elwyn Berlekamp, game theorist and coding pioneer, dies at 78
- Wall Street Journal: Math Wizard Elwyn Berlekamp Helped Bring Sharp Images From Outer Space (Paywall; download PDF version)
- Numberphile: How to Always Win at Dots and Boxes/Dots and Boxes (Extra Footage)
- Numberphile: A final game with Elwyn Berlekamp (Amazons)/Amazons (Extra Footage)
- American Go Association: Elwyn Berlekamp's Coupon Go (filmed at UC Berkeley in 2006)

President's Column *(continued from page 1)*

daily basis in areas which often require different kind of expertise than engineers usually posses.

Some other Society's members were also extremely busy in the months leading to the ISIT in Paris. We started this year with remarkably many new members at the key positions in the Society's governance, and that meant lot of learning, consulting with each other and with our predecessors, exchanging long and frequent e-mails. We got three new officers (ordinarily, there is only one), Frank Kschischang, Aylin Yener, and Wei Yu. We got a new Society's secretary, Lara Dolecek, a new conference committee chair, Vijay Kumar, a new online committee chair, Brian Kurkoski, a new external nominations committee chair, Dan Costello, a new thesis award committee chair, Christina Fragouli, a new young scholar award committee chair, Tom Fuja, a new fellows committee chair, Antonia Tulino, and new WITHITS chairs, Gireeja Ranade and Christina Lee Yu. Even people in their second year of service, like the Society's treasurer Aaron Wagner and the Newsletter editor, Salim El Rouyaheb, had entirely new types of challenges to deal with. I would like to wholeheartedly thank them all for stepping up to their respective roles when the Society needed them the most.

Most of these new appointments were made by the diligent 2018 Nominations and Appointments Committee, chaired by Alon Oriltsky. Thank you Alon and the committee. As if he did not deserve some time off after his long service to the Society, Alon continues to lead the activities related to the Shannon documentary *The Bit Player*. The movie premiered at the World Science Festival in New York City in May, and had four other screenings: at IBM, Yorktown Heights, in June, at the World Congress of Science Journalists in Lausanne and CineGlobe Festival at CERN, Geneva, in July, and at the Computer History Museum in Cupertino in August. All shows were extremely successful. You will hear about that in the next issue. For now, I recommend reading the recent review in Physics Today [2]. Yes, the physicists, again.

By the time this issue of the IT Newsletter reaches you, the IEEE annual elections will be in full swing. Please vote. Our Society is managed by the Board of Governors, and each year, one-third of the BoG gets replaced by new elected members. Please vote, and then, regardless of whether you like the outcome or not, put a serious effort to make our Society better for you and your colleagues, because leading our technical field and our technical community is too important to be left to the elected few. Is that too much to ask?

Each time there is an election of any kind, I think about the first US general elections I voted in. The year was 2004, and as usual, many were happy with the outcome and many got disappointed. The late Toni Morrison (the recipient of the 1993 Nobel Prize for literature) was among the latter. She was depressed, but eventually, inspired by a conversation with a friend, she wrote the following [3]: *This is precisely the time when artists go to work. There is no time for despair, no place for self-pity, no need for silence, no room for fear. We speak, we write, we do language. That is how civilizations heal. I know the world is bruised and bleeding, and though it is important not to ignore its pain, it is also critical to refuse to succumb to its malevolence. Like failure, chaos contains information that can lead to knowledge – even wisdom. Like art.*

References

[1] Hilbert may have said *too hard*. Many human activities (including war, by Clemenceau) are said to be too important to be left to those seemingly in charge.

[2] Toni Feder, "Review: The Bit Player, an Homage to Claude Shannon," *Physics Today*, July 2019.

[3] Toni Morrison, "No Place for Self-Pity, No Room for Fear," *The Nation, 150th Anniversary Special Issue*, April 2015.

Recent Publications

IEEE Transactions on Information Theory

Table of content for volumes 65(5), 65(6), 65(7), 65(8)

Vol. 65(5): May 2019.

Vol. 65(6): June 2019.

Vol. 65(7): July 2019.

Vol. 65(8): Aug. 2019.

Deep Learning: Mathematical Foundations and Applications to Information Science

IEEE Journal on Selected Areas in Information Theory
https://www.itsoc.org/publications/journal-on-selected-areas-in-information-theory-jsait

Call for Papers

This special issue will focus on the mathematical foundations of deep learning as well as applications across information science. Prospective authors are invited to submit original manuscripts on topics within this broad scope including, but not limited to:

- Information theoretic methods for deep learning
- Robustness for training and inference
- Understanding generalization in over-parametrized models
- Efficient and compressed model representations
- Deep generative models and inverse problems
- Large-scale efficient training of large models
- Non-convex optimization in deep learning
- Deep learning for source and channel coding.

Guest Editors

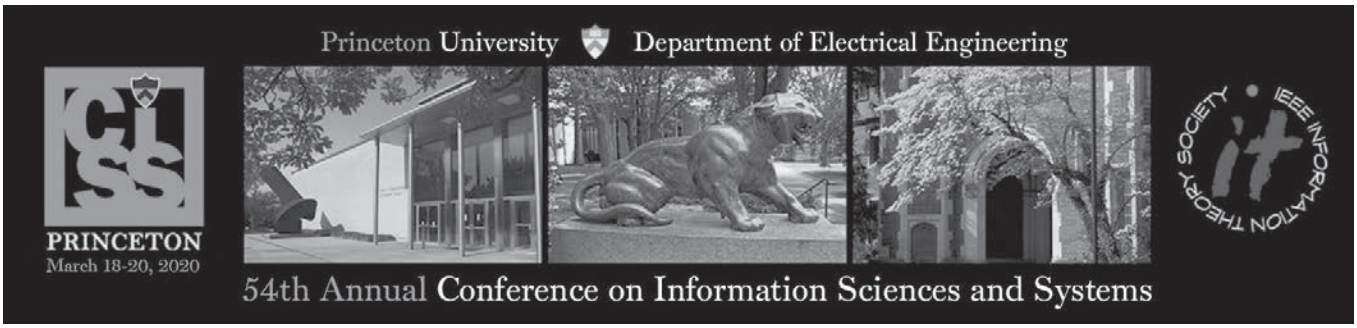
Lead Guest Editor: Alex Dimakis: dimakis@austin.utexas.edu
Richard Baraniuk: richb@rice.edu
Sewoong Oh: sewoong@cs.washington.edu
Nati Srebro: nati@ttic.edu
Rebecca Willett: willett@uchicago.edu

Submission Guidelines

Prospective authors must follow the *IEEE Journal on Selected Areas in Information Theory* guidelines regarding the manuscript and its format. For details and templates, please refer to the *IEEE Journal on Selected Areas in Information Theory* Author Information webpage. All papers should be submitted through Scholar One according to the following schedule:

Important Dates

Manuscript Due: 1 October 2019
Acceptance Notification: 15 March 2020
Final to Publisher: 5 April 2020
Expected Publication: April/May 2020



Call for Papers

54th Annual Conference on Information Sciences and Systems

March 18, 19, & 20, 2020

Princeton University - Department of Electrical Engineering

and Technical Co-sponsorship with the

IEEE Information Theory Society

Authors are invited to submit previously unpublished papers describing theoretical advances, applications, and ideas in the fields of information sciences and systems including:

·Information Theory	·Coding Theory	·Image Processing
·Communications	·Signal Processing	·Machine Learning
·Big Data Analytics	·Reinforcement Learning	·Optimization
·Statistical Inference	·Security and Privacy	·Energy Systems
·Networking	·Systems and Control	·Biological Systems

Electronic submissions of up to 6 pages (in Adobe PDF format) including 3-4 keywords must be submitted by **December 9, 2019**. Submissions should be of sufficient detail and length to permit careful reviewing. Authors will be notified of acceptance no later than **January 16, 2020**. Final manuscripts of accepted papers are to be submitted in PDF format no later than **January 30, 2020**. These are firm deadlines that will permit the distribution of electronic proceedings at the conference. Accepted papers will be allotted 20 minutes for presentation, and will be reproduced in full (up to 6 pages) in the conference proceedings. IEEE reserves the right to exclude a paper from post-conference distribution (e.g., removal from IEEE Xplore) if the paper is not presented by the author at the conference.

For more information visit us at: http://ee-ciss.princeton.edu/

CONFERENCE COORDINATOR

Lisa Lewis
Dept. of Electrical Engineering
Princeton University
Princeton, NJ 08544
Phone: (609) 258-6227
Email: ciss@princeton.edu

PROGRAM DIRECTORS

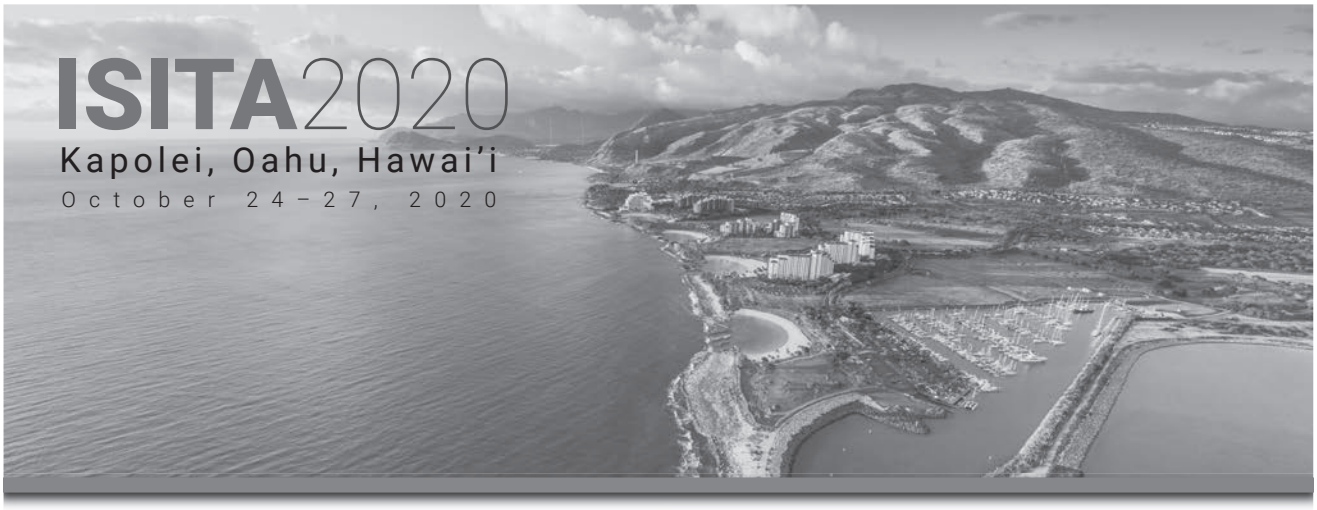
H. Vincent Poor
Yuxin Chen
Dept. of Electrical Engineering
Princeton University
Princeton, NJ 08544

IMPORTANT DATES

Paper submission deadline:
December 09, 2019

Notification of acceptance:
January 16, 2020

Final accepted manuscript due:
January 30, 2020



Symposium Committee

General Co-Chairs
Ikuo Oka Osaka City Univ.
Manabu Hagiwara Chiba Univ.
James B. Nation Univ. of Hawaii

Symposium Advisors
Toru Fujiwara Osaka Univ.
Anders Høst-Madsen Univ. of Hawaii

General Secretaries
Shigeaki Kuzuoka Wakayama Univ.
Hitoshi Tokushige Kumamoto Gakuen Univ.
Hironori Uchikawa Kioxia

Finance
Ryo Nomura Waseda Univ.
Justin Kong Univ. of Hawaii

Publicity
Brian M. Kurkoski JAIST
Akiko Manada Shonan Institute of Technology

Publications
Yu Morishima Tohoku Gakuin Univ.

Registration
Mitsugu Iwamoto
The Univ. of Electro-Communications

Local Arrangement
Takayuki Nozaki Yamaguchi Univ.
Shoko Chisaki Tokyo Univ. of Science

Technical Program Committee

TPC Co-Chairs
Hiroshi Kamabe Gifu Univ.
Navin Kashyap Indian Institute of Science

Secretary
Kenji Yasunaga Osaka Univ.

ISITA2020

October 24–27, 2020 in Kapolei, Hawai'i, USA

The **International Symposium on Information Theory and Its Applications (ISITA)** is a leading conference on information theory. Since its inception in 1990, ISITA has been a forum for interdisciplinary interaction, gathering leading researchers to discuss topics of common interest in the field of information theory. In 2020, the biennial ISITA will be held October 24–27 at Aulani, A Disney Resort & Spa in Kapolei, Hawai'i on the island of Oahu.

ISITA 2020 creates a setting for international exchange with the aloha spirit, to provide a place for individuals, especially students, to know the joy of research, and to share new results in information theory and its applications with the world.

Call for Papers

Interested authors are invited to submit papers describing novel and previously unpublished results on topics in information theory and its applications, including, but not limited to:

- Boolean Functions and Sequences
- Coding for Storage
- Coding Theory
- Communication Theory
- Computation and Complexity in Information Theory
- Cryptography and Information-Theoretic Security
- Data Privacy and Security
- Deep Learning in Information Theory
- Distributed Coding and Computation
- Estimation and Detection
- Formalization of Information Theory
- Group Testing
- Information Hiding
- Information Theory for Biology
- Information Inequalities
- Network Coding and Information Theory
- Pattern Recognition and Machine Learning
- Quantum Information and Coding Theory
- Shannon Theory
- Signal Processing
- Source Coding and Data Compression
- Sparsity and Compressed Sensing
- Statistical Inference and Learning
- Statistical Physics for Information Theory
- Statistics and Information Geometry
- Wireless Communications

Paper Submission

Authors should submit papers according to the guidelines which will later appear at:

http://isita.net

This link points to the permanent site http://www.isita.ieice.org/2020/. Accepted papers will appear in the symposium proceedings. To be published in IEEE *Xplore*, an author of an accepted paper must register and present the paper. IEEE does not guarantee inclusion in IEEE *Xplore*.

Paper submission deadline **April 2020**

Acceptance notification **June 2020**

Further information will be posted on the symposium web site as it becomes available.

Sponsor
Research Society of Information Theory and Its Applications,
Engineering Sciences Society, IEICE

Technical Co-Sponsor
IEEE Information Theory Society

Photo: Wikimedia Commons/Alakea1100

Call for Papers

2020 International Zurich Seminar on Information and Communication

February 26 – 28, 2020

The 2020 International Zurich Seminar on Information and Communication will be held at the Hotel Zürichberg in Zurich, Switzerland, from Wednesday, February 26, through Friday, February 28, 2020.

High-quality original contributions of both applied and theoretical nature in the following areas are solicited:

Wireless Communication	Optical Communication
Information Theory	Fundamental Hardware Issues
Coding Theory and its Applications	Information Theory and Statistics
Detection and Estimation	Network Information Theory and Coding
Data Storage	Cryptography and Data Security

Invited speakers will account for roughly half of the talks. In order to afford the opportunity to learn from and communicate with leading experts in areas beyond one's own specialty, no parallel sessions are anticipated. All papers should be presented with a wide audience in mind.

Papers will be reviewed on the basis of a manuscript (A4, not exceeding 5 pages) of sufficient detail to permit reasonable evaluation. Authors of accepted papers will be asked to produce a manuscript not exceeding 5 pages in A4 double-column format that will be published in the proceedings. Authors will be allowed twenty minutes for presentation.

The deadline for submission is **September 15, 2019**. Additional information will be posted at

http://www.izs.ethz.ch/

We look forward to seeing you at IZS.

Amos Lapidoth and Stefan M. Moser, Co-Chairs

Conference on Information-Theoretic Cryptography (ITC) 2020: Call for Papers

June 17–19, 2020 in Boston, MA USA

The first *Information-Theoretic Cryptography (ITC)* conference will take place on June 17-19, 2020 in Boston, MA USA. ITC is a new conference dedicated to information-theoretic aspects of cryptography, broadly defined. See the website at https://itcrypto.github.io/ for more information.

Areas of interest include, but are not restricted to:

- Randomness extraction and privacy amplification
- Secret sharing
- Secure multi-party computation
- Information theoretic proof systems
- Differential privacy
- Quantum information processing
- Oblivious data structures
- Idealized models (e.g. ,ideal channels, random oracle, generic group model)
- Bounded storage models
- Private information retrieval and locally decodable codes
- Authentication codes and non-malleable codes
- Adversarial and noisy channels
- Information-theoretic reductions
- Information-theoretic foundations of physical-layer security

Papers on all technical aspects of these and related topics are solicited for submission. Papers will be peer reviewed and accepted papers will be published in conference proceedings and presented at the conference.

The conference will have two tracks: a *publication track* and a *greatest hits* track. The publication track operates in the usual way, where authors submit their papers and the committee selects accepted papers for publication in the proceedings and presentation at the conference. The greatest hits track consists of invited talks (not published in the proceedings) that highlight the most exciting recent advances in information-theoretic cryptography. Such talks can either survey an ITC-related topic that has seen exciting developments in the last couple of years or can be devoted to a significant ITC-related result that appeared in a paper recently. This will give us the opportunity to hear about the latest big developments in information-theoretic cryptography that have appeared in different venues like FOCS/STOC,

CRYPTO/EUROCRYPT/TCC, and QIP/ISIT. The selection of speakers will be conducted by the program committee and is by invitation only. However, we solicit nominations from the community. If you would like to nominate a recent result for the greatest hits track, please send a nomination e-mail to the PC chair at itc2020chair@gmail.com. Self-nominations are discouraged.

Important Dates

- Paper Submission: Dec 16, 2019
- Greatest Hits Nomination Deadline: Jan 5, 2020.
- Acceptance Notification: March 5, 2020
- Conference: June 17-19, 2020

Conference Organization

General Chairs: Yael Tauman Kalai (MSR and MIT) and Adam Smith (BU)

Program Chair: Daniel Wichs (Northeastern and NTT Research)
itc2020chair@gmail.com

Instructions for Authors

The submission should begin with a title, followed by the names, affiliations and contact information of all authors, and a short abstract. It should contain a scholarly exposition of ideas, techniques, and results, including motivation and a clear comparison with related work. There are no other formatting requirements or page limits - it is solely up to the discretion of the authors to decide how to best present their work. It is highly recommended that authors write a good introduction, which clearly describes the main results of the paper and a high- level overview of the technical ideas.

Submissions must not substantially duplicate work that was published elsewhere, or work that any of the authors has submitted in parallel to any other journal, conference, or workshop that has proceedings. At least one author of each accepted paper is required to present the paper at the conference; presentations may be recorded and made available to the public online. Authors are strongly encouraged to post full versions of their submissions in a freely accessible online repository, such as the Cryptology ePrint archive. We encourage the authors to post such a version at the time of submission. At the minimum, we expect that authors of accepted papers will post a full version of their papers by the camera-ready deadline. Titles and abstracts of accepted papers will be made public by the PC following the notification.

Conference Calendar

DATE	CONFERENCE	LOCATION	WEB PAGE	DUE DATE
September 24–27, 2019	57th Annual Allerton Conference on Communication, Control, and Computing	Allerton, University of Illinois at Urbana-Champaign, Illinois, USA	https://allerton.csl.illinois.edu/	Passed
November 9–12, 2019	60th Annual IEEE Symposium on Foundations of Computer Science (FOCS)	Baltimore, Maryland, USA	http://focs2019.cs.jhu.edu/	Passed
November 11–14, 2019	IEEE Global Conference on Signal and Information Processing (GlobalSIP)	Shaw Center, Ottawa, Canada	http://2019.ieeeglobalsip.org/	Passed
December 9–13, 2019	IEEE Global Communications Conference (GLOBECOM)	Waikoloa, Hawaii, USA	https://globecom2019.ieee-globecom.org/	Passed
February 26–28, 2020	International Zurich Seminar on Information and Communication	Zurich, Switzerland	https://www.izs.ethz.ch/	September 15, 2019
March 18–20, 2020	54th Annual Conference on Information Sciences and Systems (CISS)	Princeton, New Jersey, USA	https://ee-ciss.princeton.edu/	December 9, 2019
October 24–27, 2020	International Symposium on Information Theory and its Applications (ISITA)	Kapolei, Hawaii, USA	http://isita.net	April, 2020

Major COMSOC conferences: http://www.comsoc.org/confs/index.html