IEEE Information Theory Society Newsletter

Vol. 62, No. 5, March 2012 Editor: Tara Javidi ISSN 1059-2362

Editorial committee: Helmut Bölcskei, Giuseppe Caire, Meir Feder, Tracey Ho, Joerg Kliewer, Anand Sarwate, and Andy Singer

President's Column

Muriél Médard

It is an honor and delight to address you for the first time as your new president. I am surrounded by a wonderful and supportive cadre of vice-presidents (Gerhard Kramer and Abbas El Gamal) as well as past presidents (Giuseppe Caire and Frank Kschischang). I would like to thank Giuseppe on behalf of the Society for his outstanding service and extend personal thanks to him for help with the transition. Giuseppe has guided the Society with warmth, sure judgment and energy. I am fortunate to have such a predecessor.

Apart from a new president, our Society also has some other transitions of officers. Andrea Goldsmith transitions off the officer roster as senior past president. Her leadership and enthusiasm have benefited the Society greatly, with such initiatives as the Student Committee, which has provided wonderful opportunities for our students with workshops, panels, as well as the very popular and well-attended Summer School. Tara Javidi replaces as Newsletter editor Tracey Ho, who finishes her term. Aylin Yener, who has performed so excellently on the Student Committee, replaces Nihar Jindal as treasurer. Many thanks to Tracey and Nihar for a job well done—the quality of the Newsletter and the health of our finances are very important to our Society. We are indebted to Tara and Aylin for agreeing to take on these important jobs.

We have an exciting year ahead of us with the upcoming IEEE review of the Society, which will take place at the TAB meeting in February (for publications only) and the TAB meeting in November (for the other activities). I have had the opportunity to observe these reviews at the TAB meeting last November and found them quite interesting. I think the exercise may prove to be a useful one in which to engage as a Society. This will be an opportunity for us to initiate a discussion about our Society's current accomplishments, operations and challenges. More importantly, it will allow us to consider collectively our vision and strategy for the future. I shall rely on the Board of Governors and on the Society as a whole for help in this important task.

In the context of seeding this conversation in our Society, I would like to offer as a topic of reflection the intellectual role our Society plays in the context not only of IEEE, but of engineering writ large. Our Society has had considerable successes in many aspects of engineering, such as progresses in physical layer communications. We have had many opportunities to reflect on those contributions. For instance, the 2004 Shannon Lecture of Bob McEliece led us through an insightful (and even musical) retrospective of the contributions of Shannon (who stood in for coding) versus those of Newton (who stood in for physics) in the context of space communications. The Society has also been at the forefront of many other developments in the physical layer space. One example among a great many is the study of MIMO systems and associated coding approaches, such as space-time codes.

However, the Society's contributions extend far beyond physical layer communications. The Society has had a central role in incubating varied domains, with considerable thematic intersection with other societies. A recent example of such a development is compressed sensing. The variety of topics of our Transactions reflects the dynamic nature of our Society's interests and the fact that our Society remains intrinsically open to new areas of investigation. Our Transactions' editor-in-chief, Helmut Bölcskei, takes great care in continually refining the composition of the editorial board to reflect the submissions, which themselves reflect the interest of our Society.

Yet, our Society may have the opportunity to contribute to engineering in even more diverse ways. The work we collectively undertake has great potential implications in communications and networking well above the physical layer, in different aspects of security, and in-network computation, which is central to cloud computing. Ubiquitous computing is predicated on ubiquitous communication and, indeed, as highlighted in the report of a recent NSF Workshop on Future Directions in Signal Processing, Information Theory, Communications and

continued on page 28

From the Editor

Tara Javidi

Dear IT Society members,

In the first issue of 2012, we have Muriel Medard's first column as IT Society President. Please join me in welcoming Muriel. With sadness, we pay tribute to our dear and brilliant friend, colleague, and mentor Jack K. Wolf who passed away in 2011. I would like to thank Roberto Padovani and Paul Siegel for sharing with us the tribute they had prepared for the National Academy of Engineering (the tribute is to appear in NAE Memorial Tribute later this year). On a happier note, we congratulate Prof. Mérouanne Debbah on his winning the prestigious 2011 SEE-IEEE Brillouin-Glavieux Prize and the newly elevated IEEE Fellows from our society. In addition to our regular contributions by Tony Ephremides and Solomon Golomb, we recap the award winning paper by Masahito Hayashi and the plenary talk by Zhanna Reznikova at ISIT 2011. We finish with two reports on the panel on New Perspectives on Information Theory at the ITW, Paraty, Brazil, and on Princeton CCI Workshop on Counting, Inference, and Optimization on Graphs prepared by Sergio Verdu and Pascal O. Vontobel, respectively.

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 jointly at the IT Society website http://www.itsoc.org/, using the quick links "Share News" and "Announce an Event". Articles and columns also can be e-mailed to me at ITsocietynewsletter@ece.ucsd.edu with a subject line that includes the words "IT newsletter." The next few deadlines are:

Issue	Deadline
June 2012	April 10, 2012
September 2012	July 10, 2012
December 2012	October 10, 2012

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. I look forward to hear your suggestions and contributions for future issues of the newsletter.

At the end of my first official editor column, I would like to thank Tracey Ho for an impeccable job as the IT newsletter editor the past three years and also for her patiently bringing me up to speed. As a person who often takes pride in her inability to avoid typos (and spills!), I was very anxious about the mechanics of the job (I even had a nightmare in which after I forgot to include the ITW panel's report, all of my IT Transactions submissions were mysteriously changed to submissions to TAC!). On the other hand, I am excited about the prospects of expanding and improving the newsletter in today's world of blogging and open access. I am especially interested to hear how the newsletter can serve and reach out to our more junior members. I have started the conversation with some of you about these questions and look forward to all of your thoughts and suggestions.

Tara Javidi

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

In Memoriam, Jack Keil Wolf 1935–2011

By Roberto Padovani And Paul H. Siegel

"For contributions to information theory, communication theory, magnetic recording, and engineering education."

JACK KEIL WOLF, a pioneer and technical leader in information theory, coding theory, communication theory, and their applications in modern information technology, died on May 12, 2011 in San Diego, California, following a battle with amyloidosis. He was 76.

Jack was born in Newark, New Jersey, on March 14, 1935. After "surviving" high school in Newark, as he would say with a smile, Jack received his B.S. in electrical engineering from the University of Pennsylvania, Philadelphia in 1956. He completed his graduate studies at Princeton, where he received the M.S.E., M.A., and Ph.D. degrees in 1957, 1958, and 1960, respectively.

Jack Keil Wolf, 1935–2011

Jack's first job was as a Lieutenant in the U.S. Air Force, working at the Rome Air Development Center in Rome, New York. At the same time, he was a part-time instructor at nearby Syracuse University which offered graduate courses at the Griffiss Air Base where Jack was stationed.

After leaving the Air Force, Jack entered a long and illustrious academic career, beginning with a position at New York University where he was a member of the Electrical Engineering department from 1963 to 1965. He then joined the Polytechnic Institute of Brooklyn in 1965 and taught there until 1973 when he joined the faculty at the University of Massachusetts, Amherst. He was a member of the Electrical and Computer Engineering department until 1984, and he served as Department Chair from 1973 to 1975.

In 1984 he joined the faculty in the Department of Electrical and Computer Engineering at the University of California, San Diego in La Jolla, California. He was appointed to an endowed chair at the newly established Center for Magnetic Recording Research. In 1993, at Jack's suggestion, the chair was renamed the Stephen O. Rice Chair in Magnetic Recording Research in honor and memory of Stephen Rice, another pioneer in communication theory and a colleague of Jack's at UC San Diego. Jack was also Vice President of Technology at Qualcomm Incorporated, which he had joined as a consultant in 1985, becoming a part-time employee in 1991. Over the course of his career Jack published more than one hundred journal papers and was granted patents on twenty-three inventions in communications and storage technology, many of which were embodied in commercial products.

Jack received many awards recognizing his technical contributions in the broad range of areas captured in the NAE citation: information theory, communication theory, magnetic recording, and engineering education. In 1975 he was co-recipient (with David Slepian) of the Information Theory Group Paper Award for the paper "Noiseless Coding for Correlated Information Sources." The main result of the paper, generally known as the "Slepian-Wolf" theorem, establishes fundamental limits on efficient distributed source coding and is considered one of the pillars of information theory. It has inspired numerous advances in both the theory and practice of data compression, with new and unforeseen applications – such as in sensor network design - emerging even today. In 1990, Jack was honored with the IEEE Communications Society E. H. Armstrong Award for "outstanding contributions over a period of years in the field of communications technology." He also shared (with Brian Marcus and Paul Siegel) the 1993 IEEE Communications Society Leonard G. Abraham Prize Paper Award for the paper "Finite-State Modulation Codes for Data Storage." In 2001 he was awarded the highest technical honor bestowed by the IEEE Information Theory Society, the Claude E. Shannon Award, and in 2007 his long record of leadership and service to the Information Theory Society was acknowledged with the Aaron D. Wyner Distinguished Service Award.

Jack's sustained contributions to the two engineering disciplines of digital communications and magnetic recording were recognized by major IEEE-level awards, namely the 1998 IEEE Koji Kobayashi Computers and Communications Technical Field Award, for "fundamental contributions to multiuser communications and applications of coding theory to magnetic data storage devices," and the 2004 IEEE Richard W. Hamming Medal, for "fundamental contributions to the theory and practice of information transmission and storage." In 2005, he was elected as a Fellow by the American Academy of Arts and Sciences.

Jack was elected to membership in the National Academy of Engineering in 1993 and the National Academy of Sciences in 2010, earning him the rare distinction of being a member of both of these academies. In 2011, he and Irwin M. Jacobs were named the winners of the Marconi Society Fellowship and Prize in recognition of "lasting scientific contributions to human progress in the field of information technology."

Jack dedicated time and energy to professional service on numerous committees of the IEEE, URSI, and NAE. He served on the Board of Governors of the IEEE Information Theory Society (then "Group") from 1970 to 1976 and from 1980 to 1986, and he was appointed President in 1974. He was also International Chairman of Committee C of URSI from 1980 to 1983. His committee work with the academies included participation in the Committee on Telecommunications Research and Development, the 2003 Nominating Committee, Electronics Engineering Section Liaison

to the NRC, Section 07 Executive and Peer Committees, member of the Committee on Tactical Battle Management, Committee on National Communications Systems Initiative, and U.S. National Committee for the International Union of Radio Science.

Jack was not only an outstanding researcher but also a dedicated and wonderful educator. He was passionate about teaching, and he had a gift for expressing in simple and clear terms even the most difficult subjects. He brought to the classroom a wealth of practical experience gained through his many years of consulting and employment in the telecommunications and storage industries. Using his unique perspective, Jack inspired his students by successfully linking elegant theory with exciting technological applications. In 2000, Jack's excellence in teaching was recognized with the UCSD Distinguished Teaching Award.

Jack maintained a close relationship with his alma mater, the University of Pennsylvania. In fact, studying at Penn was somewhat of a family tradition: seventeen other members of Jack's extended family – including his father, two uncles, numerous cousins, and daughter Sarah – received degrees from Penn, and a grandson and granddaughter are carrying the torch for the next generation. Jack and his daughter have also made philanthropy at Penn a tradition: a number of endowed scholarships and student awards bear the Wolf family name, and two laboratories are named in honor of Jack. In 2006, Jack received the D. Robert Yarnall Award from the University of Pennsylvania Engineering School, an award presented annually to a distinguished member of Penn Engineering's alumni for outstanding contributions to society in the field of engineering or technology.

Jack is deeply missed by his family, friends, and colleagues, including his many students, past and present, affectionately known as the "Wolf Pack." What Jack brought to the classroom and research advising was much more than his gift and passion for teaching: he inspired generations of students to excel, to work hard and with integrity, and most of all to have fun in the process. Jack will be remembered for the friendship and support he and his wife Toby offered so freely, his smile and sense of humor, his vision and wisdom, his words of encouragement, and his contagious optimism. He was a generous, thoughtful, and unpretentious man, an exceptional human being dedicated to bettering our world through progress in engineering.

A devoted husband, father, and grandfather, Jack is survived by his wife Toby, his children, Joe, Jay, Jill, Sarah and her husband Charles, and his grandchildren, Rachel, David, Rebecca, Aaron, and Julia.

This article will be published in "Memorial Tributes: National Academy of Engineering, Volume 16."

IT Society Member Wins IEEE Medal

Mérouane Debbah, professor at Supélec and head of the Alcatel-Lucent chair on Flexible Radio, received on the 7th of December 2011 the joint IEEE and SEE Brillouin-Glavieux 2011 award. The ceremony took place in Paris at the SEE (Société de l'Electricité, de l'Electronique et des Technologies de l'Information et de la Communication) headquarters in the presence of Mme Glavieux and Martin Bastiaans, director of the region 8 of IEEE. The award "recognizes the scientific excellence of Prof. Debbah's work and contributions in the field of Information Theory and its applications to Wireless Communications". Mérouane Debbah is an expert in the field of Random Matrices and Game Theory and made numerous contributions in their applications to Wireless Communications. Established by the SEE and IEEE, the prize awards each year a young researcher (under 40) who made important contributions in the field of science and innovation in the area of Telecommunications.

Mérouane Debbah entered the Ecole Normale Supérieure de Cachan (France) in 1996 where he received his M.Sc and Ph.D. degrees respectively. He worked for Motorola Labs (Saclay, France) from 1999–2002 and the Vienna Research Center for Telecommunications (Vienna, Austria) from 2002–2003. He then joined the Mobile Communications department of the Institut Eurecom (Sophia Antipolis, France) as an Assistant Professor. Since 2007, he is a Full Professor at Supelec (Gif-sur-Yvette, France), holder of the Alcatel-Lucent Chair on Flexible Radio. Mérouane Debbah is the recipient of the "Mario Boella" prize award in 2005, the 2007 General Symposium IEEE GLOBECOM best paper award, the Wi-Opt 2009 best paper award, the 2010 Newcom++ best paper award as well as the Valuetools 2007, Valuetools 2008 and CrownCom2009 best student paper awards. He is a WWRF fellow.

2012 Newly Elevated Fellows

Erdal Arikan
Bilkent University Ankara

for contributions to coding theory

Martin Bossert
Ulm University

for contributions to reliable data transmission including code constructions and soft decision decoding

Roger SK Cheng
Hong Kong University of Science & Technology

for contributions to multiuser communications in wireless systems

Stefano Galli
ASSIA, Inc.

for contributions to theory, practice, and standardization of power line communication networks

Ryuji Kohno
Yokohama National University

for contributions to spread spectrum and ultra wideband technologies and applications

Adam Krzyzak
Concordia University

for contributions to nonparametric algorithms and classification systems for machine learning

Soung Chang Liew
The Chinese University of Hong Kong

for contributions to wireless communications and networking

Ranjan K. Mallik
Indian Institute of Technology Delhi

for contributions to channel characterization in wireless communication systems

Eric Lawrence Miller
Tufts University

for contributions to inverse problems and physics-based signal and image processing

Eytan Modiano
Massachusetts Institute of Technology

for contributions to cross-layer resource allocation algorithms for wireless, satellite, and optical networks

Jong-Seon No
Seoul National University

for contributions to sequences and cyclic difference sets for communications algorithms

Jean-Christophe Pesquet
University Paris-Est

for contributions to statistical methods for signal recovery

Konstantinos N. Plataniotis
University of Toronto

for contributions to the theory and application of statistical adaptive learning

Wonjong Rhee
ASSIA, Inc.

for leadership in dynamic spectrum management systems

Akbar M. Sayeed
University of Wisconsin-Madison

for contributions to statistical signal modeling for wireless communication and sensor networks

Ljubisa Stankovic
University of Montenegro

for contributions to time-frequency signal analysis

Emre I. Telatar
Ecole Polytechnique Federale de Lausanne

for contributions to information theory and coding

Bane Vasic
University of Arizona

for contributions to coding theory and its applications in data storage systems and optical communications

Jiangtao Wen
Tsinghua University

for contributions to multimedia communication technology and standards

Guu-Chang Yang
National Chung Hsing University

for contributions to optical code division multiple access

Junshan Zhang
Arizona State University

for contributions to cross-layer optimization of wireless networks

The Historian's Column

Anthony Ephremides

It is crisis-time in the world today: crisis in the financial world, turmoil in the world economy, disorder in Society, and soul-searching in Science. Periodically such phenomena of agony, doubt, concern, and pessimism tend to emerge and dominate and then subside or go into a sleep-mode. Over the years, we have observed such a pattern within the narrow domain of our field as well. Since the days of optimistic abandon that followed the birth of Information Theory, there were many periods of retrenchment, disappointment, and direct confrontation and challenge. As we have reminisced in the past, the field of Information Theory has experienced many births and deaths, which has confirmed its inherent vitality. Just as we reach vertical walls of technical difficulty, there is always a new ray of light that penetrates through and allows us to move forward.

I was having these thoughts recently as I was reading about the debt crisis in most countries of the world and, at the same time, contemplating negative reviews that I received for a proposal that I thought was full of merit, imagination, and, as the cliché goes, "high-risk, high-reward" ideas. Is the world of Science and Technology experiencing a similar crisis? Has our reach exceeded our grasp? Are we building on an unstable foundation an unsustainable superstructure?

There is increasing evidence that the answer may be affirmative. The number of journals is increasing by the day, while the quality control they exercise is clearly diminishing. The number of conferences is getting to be a joke. Every day we see announcements for the "First International Symposium on…you name it". The number of submitted papers is exploding. The number of eligible, willing, and conscientious reviewers is decreasing. The number of new "initiatives" for funding is proliferating. The number of panels to judge and decide on fund disbursement is increasing while their quality is decreasing. Feelings of frustration abound. It is getting harder for the graduates from the best Institutions to find jobs. What used to be the job of accomplished researchers is more and more becoming delegated to junior and inexperienced ones, including beginning PhD students. What is happening? Why? What can be done?

These observations and questions parallel in many respects what we see in the world at large today. Philosophers, economists, and politicians are trying to make ends and tails of it but so far there does not seem to be any consensus. The world continues to function and, by hook or by crook, things get done and we limp forward.

In this generalized mess, are there islands of serenity, excellence, and contentment? I am sure there are. For example, in the hearts and minds of many of us there must be feelings and thoughts of satisfaction and optimism; but what about large groups, communities, and organizations? How about Universities, technical Societies, and organizations?

I cannot answer that. Perhaps there are. But I would like to offer some thoughts on how to "cope", if you will, with the stalemate that many of us perceive to be developing, at least within the confines of our small (and select) community.

First, and foremost, we need to have a sense of humor. There is no better antidote to depression than a light mood and a realization that the best defense is to laugh at adversity. Try to imagine the anonymous, unscrupulous reviewer who trashed your best work and point your finger to him/her and burst into laughter. Curse him with a "may all your theorems be wrong, you fool!" Then, do not give up. As an unsavory character in opera says: "Insistiamo!", i.e. persist ! Not by resending the same paper or proposal elsewhere, but by going back to the drawing board and generating new ideas. Finally, look back for inspiration. Doomed are those who forget their heritage. There are examples of actions and ideas of our predecessors that should continue to inspire us. What would Shannon have done if his paper on the binary symmetric channels had been rejected? Someone might say "he would have gone back to juggling"; but I would say, he would have come back with more and better explanations for the narrow- minded audiences.

From personal experience, let me relay to you an incident early in my career. In 1987 I was sitting on the Board of Governors of the Control Systems Society and at the same time I happened to be the President of the Information Theory Society. At a meeting of the Control Systems Society BoG, the then President of it, Jim Melsa made a presentation full of agony about the future of the Control Systems Society. For added drama he decided to make his case of alerting the society members to confront the looming disaster by saying "Let us not allow what happened to the Information Theory Society (then Group), happen to us; let us face it: Information theory is DEAD! ". You can imagine my position. I am sitting there as the president of a dead society. What an insult! What an abomination! My adrenaline was stirred and, as I do not need much to fire me up, I got up, interrupted and said proudly: "I take exception to the President's remark about the demise of Information Theory". There was silence in the room! And that loud protest, since that cold winter day of 1987, ensured the survival and prosperity of our Society in the years that followed!

This is how I single-handedly reversed the death announcement and changed the course of History. Let this be an example of how to confront the looming crises. Simply refuse to accept them. Then History will follow its course. Or, on a more serious note, "stay the course"! Then History will follow yours.

Second Order Analysis Based on Information Spectrum

Masahito Hayashi

Abstract – In this letter, we explain the importance of the 2nd-order asymptotic theory for channel coding and its mathematical structure via the information spectrum method, which provides an unified viewpoint for the 2nd-order asymptotic theory in information theory. Further, we treat applications to quantum cryptography, folklore in source coding, ans security analysis.

Key words – 2nd-order asymptotic theory, information spectrum method, channel coding, central limit theorem

I. Introduction

The asymptotic theory is one of the most important topics in information theory. This is because the optimal performance can be described by the use of entropy or the mutual information in the asymptotic regime while it does not have a simple description in the finite length regime. Hence, the asymptotic theory has been studied as one of the main topics in information theory.

Fortunately, the asymptotic theory has been successfully applied to various kinds of problems in information theory. These asymptotic theory have treated the case when the bit length L_n behaves as $R_1 n$ with the block length n. In such a case, we focus on the minimum of the limit of the average error probability for the coding rate R_1. The optimal rate R_1 with the asymptotic zero error probability is called the optimal coding rate, and has been one of the main topic in this area.

However, there arises the question: it is too simplified to describe the bit length L_n as $R_1 n$. Hence, we adopt the description expanding L_n to $R_1 n + R_2 \sqrt{n}$, in which, the asymptotic minimum average error probability is treated based on the the second order coefficient R_2 as well as the first order coefficient R_1. This type asymptotic theory is called the second order asymptotic theory while the asymptotic theory based only on the first order coefficient R_1 is called the first order asymptotic theory. In general, the asymptotic theory based up to the k-th order coefficient is called the k-th order asymptotic theory. It is usual to increase the degree of the order in many research areas. For example, in statistics, the asymptotic treatment of the mean square error is one of the main topics, and its asymptotic expansion up to the second order is called the second order asymptotic theory [1].

One might consider that the second order asymptotic theory is a relief work for researchers that can find no suitable topic among the first order asymptotic theory. However, this is not true because there exists a big limitation in the first order asymptotic theory in information theory and it can be resolved by the second order asymptotic theory.

The first order asymptotic theory guarantees the existence of the reliable code with the optimum rate. Hence, we can concentrate to optimize our code among codes within codes with the optimum rate. However, as is described in latter, there is no suitable answer for the question what performance is available for a given rate R_1 and a given finite block size n even with approximation. This question can be answered by the second order asymptotic theory approximately. This is a point different from the second order asymptotic theory in statistics. The effect of the second order asymptotic theory in statistics is improving the accuracy of the mean square error. The second order asymptotic theory in information theory can resolve the problem that cannot be resolved by the first order asymptotic theory at all. That is, it can provide an approximation value for the optimum average error probability for a fixed bit length L_n and and a given finite block size n. Indeed, in order to resolve this problem, we require the second order expansion at least.

From the mathematical viewpoint, the first order asymptotic theory in information theory corresponds to the law of large numbers, and the second order asymptotic theory in information theory does to the central limit theorem. The consistency for estimator in statistics corresponds to the first order asymptotic theory in information theory because it is essentially based on the law of large numbers. Then, the first order asymptotic theory in statistics corresponds to the second order asymptotic theory in information theory because it is essentially based on the central limit theorem. As is described latter, the relations with the law of large numbers and the central limit theorem can be clarified by the treatment via the method of information spectrum.

On the other hand, the second order asymptotic theory in statistics treats the more detail structure based on the curvature [2]. So, we can conclude that the order in information theory does not match that in statistics. This is because information theory focuses on a value different from the value focused in statistics while the order of asymptotic theory is determined by the asymptotic expansion of the focused value. In particular, the second order asymptotics in statistics treats so detailed structure that the optimization depends on the existence of the bias correction. However, the second order asymptotics in information theory does not treat so detailed structure. This fact means that the meaning of the optimum value in the second order asymptotic theory in information theory is much clear than that in statistics. Hence, it had better be discussed earlier. However, there exist only second order studies for variable-length coding [3], [4], [5], [6], [7] and for channel coding by Strassen [8] before the author started the studies for the second order asymptotic theory. All of them has no unified treatment concerning the second order asymptotic theory.

The author treated the second order asymptotic theory [9], [10] for the channel coding, fixed-length source coding, and uniform random number generation from a unified viewpoint by employing the method of information spectrum established by Han [11]. It is valuable to an approximation of the optimal performance with a fixed size because the approximation clarifies the possibility for further improvement of the realized code. In the following, we briefly explain the method of information spectrum. As is illustrated in Fig. 1, this method treats the problem via two steps.

In the first step, the optimum asymptotic performance is derived for a given information source or a given channel without any condition, e.g., the independent and identical distributed (i.i.d.) condition or the Markovian condition. The optimum asymptotic performance is represented by a limiting value defined by the logarithmic likelihood or the logarithmic likelihood ratio. Such quantities are called information spectrum quantities. Since these values represent no concrete values, only this step cannot resolve the problem.

In the second step, we calculate the information spectrum quantities concretely in the respective cases, e.g., the i.i.d. case or the Markovian case. Fortunately, it is easy to calculate these values in the i.i.d. case and the Markovian case. Only the first step depends on the type of the information process. The second step depends on the type of information sources or channels, but does not depend on the type of the information process. Especially, when variable types of information process give the same the information spectrum quantity as their optimal performance, the existing calculation can be recycled in the second step.

On the other hand, the first step for the first order asymptotics can be directly applied that for the second order asymptotics. That is, as soon as this problem is mathematically formulated, the solution has been already given [10]. Such a story is too convenient for researchers. Therefore, only the second step is required for the second order asymptotic theory. It can be also resolved only by application of the central limit theorem except for channel coding. only the impossibility part (the converse part) for channel coding cannot be resolved in the above simple way. Hence, the second order asymptotic theory can be visibly discussed via the method of information spectrum. Due to this kind of advantage, recently, Nomura et al. [12] treated the second order asymptotics for intrinsic randomness by the same way. Further, the second order asymptotics can be applied to security analysis for quantum key distribution, folklore in source coding, and the limitation of the traditional security analysis via the error correction of the dummy variable.

The following is the organization of the remaining parts. In Section II, in order to explain why it is impossible to approximately evaluate the optima performance with the finite size regime, we first review the law of large numbers and the central limit theorem. In Section III, we treat the second order asymptotics for channel coding with the relation to the law of large numbers and the central limit theorem. In Section IV, the extension of the additive white Gaussian noise (AWGN) case is treated. Moreover, in Section V, we explain the method of information spectrum and how applied it to the second order asymptotic theory. In Section VI, we review how the second order asymptotics is applied to the quantum key distribution. In Section VII, we treat folklore in source coding. Finally, in Section VIII, we discuss the limitation of the security analysis based on the error correction of the dummy variable by employing the second order asymptotics.

II. Law of Large Numbers and Central Limit Theorem

The law of large numbers and the central limit theorem are most fundamental topics in an elementary course of probability and statistics. Let X be the random variable and $E(X)$ be its expectation. Assume that the independent random variables $X_1, \ldots, X_n$ subject to the same distribution as the random variable X.

Then, the sample mean $X^n := ((X_1 + \cdots + X_n)/n)$ satisfies the law of large numbers as follows:

Theorem 1 (Law of large numbers): For any real number $\epsilon > 0$, we obtain the following relation

$$\Pr\{|X^n - E(X)| > \epsilon\} \to 0. \tag{1}$$

However, the above theorem cannot answer the question how close to zero the LHS is for a given finite n and a given real number $\epsilon > 0$. For example, it is impossible to give its approximation even with three significant figures. Instead, the order of the LHS cannot be answered with n and $\epsilon > 0$.

One reason is that the limit of this limiting formula is zero. Due to this reason, we cannot give its approximation even with three significant figures. Another reason is that the convergence is not uniform when $\epsilon > 0$ is close to zero. Since interesting region is the neighborhood of zero, the limit might not work properly even if the limit is not zero. Hence, the law of large numbers does not provide no more than that the LHS becomes smaller when the integer n becomes larger. Therefore, this theorem cannot provide any information for a finite n and a given $\epsilon > 0$, which is a crucial defect of this theorem.

This is because this theorem treats only the coarse asymptotic behavior for the sample mean X^n. For example, while a microbe has internal structure, it looks only one point when we watch it by bare eye. This is because our bare eye can watch only the coarse scale. In order to watch the detail structure, we have to use the microscope that accommodates the scale of the microbe.

Similarly, in order to treat the detail of the asymptotic behavior of the sample mean X^n, we have to focus on the expectation $E(X)$ and amplify the neighborhood of $E(X)$. The central limit theorem is the theorem revolving the above problem by this kind of amplification. Let $V(X)$ be the variance of X. Using the cumulative distribution function of the standard normal distribution $\Phi(a) := \int_{-\infty}^{a} (1/\sqrt{2\pi})\, e^{-x^2/2}\, dx$, we can give the central limit theorem as follows.

Theorem 2: For any $\epsilon > 0$, the following relation holds.

$$\Pr\left\{\left|\frac{\sqrt{n}\,(X^n - E(X))}{\sqrt{V(X)}}\right| > \epsilon\right\} \to \Phi(-\epsilon) + 1 - \Phi(\epsilon) \tag{2}$$

For an approximation of the LHS of (1), it is sufficient to deform the LHS of (2) by resetting the parameter $\epsilon > 0$. Then, the RHS of (2) can be regarded as its approximation.

Indeed, since the RHS of (2) is not zero, it can provides an desired approximation. Further, the convergence of this theorem is uniform when $\epsilon > 0$ is close to zero. Hence, we can expect that the RHS of (2) is sufficiently close to the true value of the LHS of (2) with a finite n. In the case of binomial distribution, the RHS of (2) might coincide with the LHS of (2) with three significant figures. except for the extremal case when $\epsilon > 0$ is sufficiently smaller than 1 and n is around 100. The central limit theorem properly works as a method approximating an interesting probability in this way.

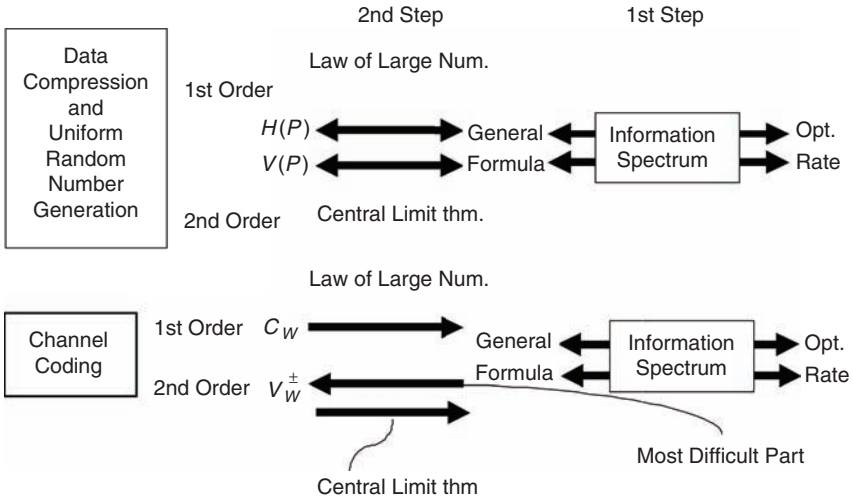


Fig. 1. Structure of the first and second asymptotic theory via the method of information spectrum.

III. Second Order Asymptotics for Channel Coding

When we send the message via n use of the channel $W(y|x)$ from the sets of alphabets to be sent $\mathcal{X}$ to the sets of alphabets to be received y, we focus on the average output distribution $W_P(y) := \sum_x P(x)W(y|x)$ for a given input distribution P. Then, the mutual information is written by $I(P,W) := \sum_x P(x)\sum_y W(y|x)\log\left(W(y|x)/W_P(y)\right)$. In this case, as is shown as the channel coding theorem, the asymptotic optimal transmission rate is given by the capacity $C(W) := \max_P I(P,W)$.

Let $\mathcal{M}_n := \{1,\ldots,M_n\}$ be the set of messages to be transmitted. Then, the encoder ψ_n is given as a map from $\mathcal{M}_n$ to $\mathcal{X}^n$. The decoder ϕ_n is given as a map from $\mathcal{Y}^n$ to $\mathcal{M}_n$. The average error probability is written as

$$\epsilon_n(\psi_n,\phi_n) := 1 - \frac{1}{|\psi_n|}\sum_{i=1}^{|\psi_n|}\sum_{y\in\phi_n^{-1}(i)} W^n(y|\psi_n(i)), \tag{3}$$

where $|\psi_n|$ is the size of the message set $\mathcal{M}_n$. The minimum average error probability for a fixed transmission bit L is

$$C_n(L|W) := \min_{(\psi_n,\phi_n)} \{\epsilon_n(\psi_n,\phi_n)|\log|\psi_n| \ge L\}. \tag{4}$$

It is known that this value is calculated as

$$\lim_{n\to\infty} C_n(nR_1|W) = \begin{cases} 1 & \text{if } R_1 > C(W) \\ 0 & \text{if } R_1 < C(W). \end{cases} \tag{5}$$

As is shown in (5), the minimum error probability is changed from 0 to 1 when the transmission rate R_1 is $\max_P I(P,W)$. This property is called the strong converse.

However, a more interested value from the practical viewpoint is the value $C_n(nR_1|W)$ with a finite number n. Unfortunately, the above limit formula (5) does not give an approximation for the value $C_n(nR_1|W)$ with finite number n. This is because (5) gives the value zero when $R_1 < C(W)$ As another reason, the convergence in (5) is not uniform concerning R_1. Due to these reasons, the limit formula (5) is no more an approximation for the value $C_n(nR_1|W)$ with any finite number n than the law of large number is an approximation for the probability.

In order to resolve this problem, we treat the transmission rate up to the second order $\sqrt{n}$. While we surpass the condition $\log|\psi_n| \le R_1 n$ in the above discussion, we surpass a more precise condition $\log|\phi_n| \le R_1 n + R_2\sqrt{n}$ in the following. The minimum average error probability under this condition is given by $\lim_{n\to\infty} C_n(R_1 n + R_2\sqrt{n}|W)$. When the channel is given as an additive channel, i.e., there exists a probability distribution Q such that $W(y|x) = Q(y-x)$, this value can be calculated by using the values $H(Q) := -\sum_x Q(x)\log Q(x)$ and $V(Q) := \sum_x Q(x)(-\log Q(x) - H(Q))^2$ in the following way:

$$\lim_{n\to\infty} C_n(R_1 n + R_2\sqrt{n}|W) = \begin{cases} 0 & \text{if } R_1 < \log d - H(Q) \\ \Phi(R_2/\sqrt{V(Q)}) & \text{if } R_1 = \log d - H(Q) \\ 1 & \text{if } R_1 > \log d - H(Q). \end{cases} \tag{6}$$

For a general channel W, the formula is more complicated because a distribution P satisfying $I(P,W) = \max_{P'} I(P',W)$ is not unique in general. Then, we treat the following two quantities:

$$V^+(W) := \max_P \sum_x P(x)\sum_y W_x(y)\left(\log\frac{W_x(y)}{W_P(y)} - D(W_x\|W_P)\right)^2$$

$$V^-(W) := \min_P \sum_x P(x)\sum_y W_x(y)\left(\log\frac{W_x(y)}{W_P(y)} - D(W_x\|W_P)\right)^2.$$

where $W_x(y) := W(y|x)$ and the KL-divergence is defined as $D(P\|Q) := \sum_x P(x)\log\left(P(x)/Q(x)\right)$ and max and min denote the maximization and the minimization among distributions P satisfying $I(P,W) = C(W)$. Then, $\lim_{n\to\infty} C_n(R_1 n + R_2\sqrt{n}|W)$ is calculated by

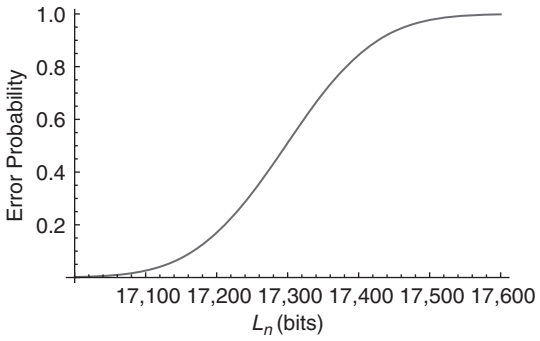


Fig. 2. The approximate average error probability $\Phi((L_n - (n/2)\log(1+(P/N)))/\sqrt{nV_{N,P}})$ with $n = 10000$, $P/N = 10$.

$$\lim_{n\to\infty} C_n(R_1 n + R_2\sqrt{n}\,|W) = \begin{cases} 0 & \text{if } R_1 < C(W) \\ \Phi(R_2/\sqrt{V^-(W)}) & \text{if } R_1 = C(W), R_2 \le 0 \\ \Phi(R_2/\sqrt{V^+(W)}) & \text{if } R_1 = C(W), R_2 > 0 \\ 1 & \text{if } R_1 C > (W). \end{cases} \quad (7)$$

The same calculation is possible for the Gaussian case and the case with energy constraint [9].

These fact can be shown via the method of information spectrum. The limiting behavior of the second order asymptotics of the channel coding is quite similar to that of the random variable in the central limit theorem. However, the above relation is not outward but roots in a deeper connection. As is explained in the next section, the central limit theorem plays an important role in the derivation of the limit formula (7).

Further, when we fix R_1 to be the capacity $C(W)$, the quantity $\lim_{n\to\infty} C_n(R_1 n + R_2\sqrt{n}\,|P^X)$ is not zero. Further, this convergence is uniform for any compact region. Therefore, the quantity $\lim_{n\to\infty} C_n(R_1 n + R_2\sqrt{n}\,|P^X)$ works for an approximation for $C_n(R_1 n + R_2\sqrt{n}\,|P^X)$ with a sufficiently large n. Therefore, the second order asymptotics can provide an approximation for the minimum average error probability with a finite size n although the first order asymptotics cannot provide it.

The strong converse is considered as an important property in information theory (including quantum information theory), Hence, it is treated as an important topic to show this property. However, even if the strong converse for a topic in information theory is shown, there remains another important problem for the asymptotic behavior of this topic. In fact, the minimum error probability for a fixed rate cannot be approximated by its limit when the strong converse holds. Therefore, we need the second order asymptotic theory for approximation of the minimum error probability for a fixed rate with a finite length size n when the strong converse holds. From a technical viewpoint, the second order asymptotics is more difficult than the strong converse. Then, the second order asymptotics should be treated after solving the strong converse.

Since the minimum error probability $C_n(R_1 n|W)$ goes to zero exponentially when R_1 is less than the capacity $C(W)$, the traditional researches treated the finite-size effect for $C_n(R_1 n|W)$ by treating the exponential decreasing rate. The most famous approach of this direction is Gallagar's bound [13]. Gallagar's bound [13] yields an upper bound of $\lim_{n\to\infty} C_n(C(W)n + R_2\sqrt{n}\,|W)$, the upper bound is almost twice of the value $\lim_{n\to\infty} C_n(C(W)n + R_2\sqrt{n}\,|W)$. Hence, the traditional approach of exponent is not necessarily the best method for evaluating the minimum average error $C_n(L|W)$. Since the second order asymptotic theory provides the limit $\lim_{n\to\infty} C_n(C(W)n + R_2\sqrt{n}\,|W)$, it can gives an approximation value closer to the minimum average error $C_n(L|W)$. Therefore, we can conclude that the second order asymptotic theory has the same importance as the central limit theorem in statistics.

IV. Additive White Noise Gaussian Case

Now, we treat the generalization of the formula (7) to the additive white noise Gaussian case. Assume that the additive noise X is subject to the Gaussian distribution with the variance N. We also assume the so-called "power" constraint, requiring that for a codeword $(x_1, x_2, \ldots, x_k)$ transmitted through the channel, we have: $(1/n)\sum_{i=1}^n x_i^2 \le P$. Under this constraint, Similar to (4), we define the minimum average error probability $C_n(L|N,P)$ with the block length n and the transmitted information length L. Then, the capacity is known to be $(1/2)\log(1+(P/N))$. Using the quantity $V_{N,P} := ((P^2/N^2) + 2(P/N))/(2(1+(P/N)))^2$, the formula (7) can be generalized to

$$\lim_{n\to\infty} C_n(R_1 n + R_2\sqrt{n}\,|N,P) = \begin{cases} 0 & \text{if } R_1 < \frac{1}{2}\log(1+\frac{P}{N}) \\ \Phi(R_2/\sqrt{V_{N,P}}) & \text{if } R_1 = \frac{1}{2}\log(1+\frac{P}{N}) \\ 1 & \text{if } R_1 > \frac{1}{2}\log(1+\frac{P}{N}). \end{cases} \quad (8)$$

Hence, using the formula (8) with $R_1 = (1/2)\log(1+(P/N))$, we can approximate the minimum average error probability $C_n(L_n|N,P)$ to be

$$\Phi((L_n - \frac{n}{2}\log(1+\frac{P}{N}))/\sqrt{nV_{N,P}}) \quad (9)$$

because $C_n(L_n|N,P)$ can be approximated by $\Phi(R_2/\sqrt{V_{N,P}})$ with R_2 satisfying $L_n = (n/2)\log(1+(P/N)) + R_2\sqrt{n}$.

V. The Method of Information Spectrum

Next, we focus on the method of information spectrum, which enables us to treat the second order asymptotics in a parallel way to the first order asymptotics. The method of information spectrum is a general theory that can be applied to a general sequence of channels $\bar{W} := \{W^n(y|x)\}_{n=1}^\infty$, whose input alphabet set and output alphabet set are given as $\mathcal{X}_n$ and $\mathcal{Y}_n$.

Although it is usual to assume the memoryless condition, the Markov condition, or the stationary condition, the following discussion does not assume any condition for a sequence of channels. Here, for any sequence of distribution on the input alphabet set $\bar{P} := \{P^n\}$, we define the following quantities:

$$\overline{P}_I(R_1|\bar{P},\bar{W}) := \limsup_{n\to\infty} \sum_{x\in\mathcal{X}_n} P^n(x) W_x^n \left\{\log\frac{W^n(y|x)}{W_{P^n}^n(y)} \le R_1 n\right\}$$

$$\underline{P}_I(R_1|\bar{P},\bar{W}) := \liminf_{n\to\infty} \sum_{x\in\mathcal{X}_n} P^n(x) W_x^n \left\{\log\frac{W^n(y|x)}{W_{P^n}^n(y)} \le R_1 n\right\},$$

where $W_x^n(y) := W^n(y|x)$ and $W_{P^n}^n(y) := \sum_x P^n(x) W^n(y|x)$.

When we fix the size $|\psi_n|$ for a respective integer n, the minimum average error probability is given as

$$C_n(L|W^n) := \min_{(\psi_n,\phi_n)} \{\epsilon_n(\psi_n,\phi_n) | \log|\phi_n| \ge L\}. \tag{10}$$

Then, we can obtain the following formula:

$$\lim_{\gamma\to+0} \inf_{\bar{P}} \underline{P}_I(R_1-\gamma|\bar{P},\bar{W}) \le \liminf_{n\to\infty} C_n(R_1 n|W^n) \le \lim_{\gamma\to+0} \inf_{\bar{P}} \underline{P}_I(R_1+\gamma|\bar{P},\bar{W}) \tag{11}$$

$$\lim_{\gamma\to+0} \inf_{\bar{P}} \overline{P}_I(R_1-\gamma|\bar{P},\bar{W}) \le \limsup_{n\to\infty} C_n(R_1 n|W^n) \le \lim_{\gamma\to+0} \inf_{\bar{P}} \overline{P}_I(R_1+\gamma|\bar{P},\bar{W}). \tag{12}$$

In particular, when four quantities $\lim_{\gamma\to+0}\inf_{\bar{P}}\underline{P}_I(R_1-\gamma|\bar{P},\bar{W})$, $\lim_{\gamma\to+0}\inf_{\bar{P}}\underline{P}_I(R_1+\gamma|\bar{P},\bar{W})$, $\lim_{\gamma\to+0}\inf_{\bar{P}}\overline{P}_I(R_1-\gamma|\bar{P},\bar{W})$, $\lim_{\gamma\to+0}\inf_{\bar{P}}\overline{P}_I(R_1+\gamma|\bar{P},\bar{W})$ coincide, the limit $\lim_{n\to\infty} C_n(R_1 n|W^n)$ exits and coincides with them, i.e.,

$$\lim_{n\to\infty} C_n(R_1 n|W^n) = \lim_{\gamma\to+0} \inf_{\bar{P}} \underline{P}_I(R_1-\gamma|\bar{P},\bar{W}). \tag{13}$$

Therefore, we can evaluate the limits $\liminf_{n\to\infty} C_n(R_1 n|W^n)$ and $\limsup_{n\to\infty} C_n(R_1 n|W^n)$ by using the above four quantities for any sequence of channels. These formulas do not provide explicit bounds for these limits. Hence, their meaning is not as clear as entropy.

However, in the case of discrete memoryless, as is shown via slightly complicated calculation, the above four quantities coincide and are written by using the capacity $C(W)$ as follows.

$$\inf_{\bar{P}} \lim_{\gamma\to+0} \underline{P}_I(R_1+\gamma|\bar{P},\bar{W}) = \begin{cases} 1 & \text{if } R_1 > C(W) \\ 0 & \text{if } R_1 < C(W). \end{cases} \tag{14}$$

Due to this fact, the minimum average error probability converges to zero when the transmission rate is less than the capacity $C(W)$, i.e., we obtain (14).

When the set $\mathcal{X}_n$ equals $\mathcal{Y}_n$ and has a module structure and the channel $W^n(y|x)$ is written by using a distribution Q^n on the module $\mathcal{X}_n$:

$$W^n(y|x) = Q^n(y-x), \tag{15}$$

the channel $W^n(y|x)$ is called additive. When $|\mathcal{X}| = d^n$, we can show that

$$\overline{P}_I(R_1|\bar{P},\bar{W}) \ge 1 - \underline{P}_H(\log d - R_1|\bar{Q}),$$

where

$$\underline{P}_H(R_1|\bar{Q}) := \liminf_{n\to\infty} Q^n\left\{\frac{-1}{n}\log Q^n(x) \le R_1\right\}. \tag{16}$$

Since the equality holds when $\bar{P}$ is the uniform distribution, we obtain

$$\inf_{\bar{P}} \overline{P}_I(R_1|\bar{P},\bar{W}) = 1 - \underline{P}_H(\log d - R_1|\bar{Q}).$$

Similarly, we obtain

$$\inf_{\bar{P}} \underline{P}_I(R_1|\bar{P},\bar{W}) = 1 - \overline{P}_H(\log d - R_1|\bar{Q}),$$

where

$$\overline{P}_H(R_1|\bar{Q}) := \limsup_{n\to\infty} Q^n\left\{\frac{-1}{n}\log Q^n(x) \le R_1\right\}. \tag{17}$$

Therefore, when Q^n is the independent and identical distribution of the distribution Q, we obtain

$$\overline{P}_H(R_1|\bar{Q}) = \underline{P}_H(R_1|\bar{Q}) = \begin{cases} 1 & \text{if } R_1 > H(Q) \\ 0 & \text{if } R_1 < H(Q), \end{cases} \tag{18}$$

which implies (5).

In the following, we prove (11). For this purpose, we prepare the following two lemmas.

Lemma 1 (*Verdú-Han* [14]): For any channel $W^n(y|x)$, any distribution $P^n(x)$, any integer M_n, and any real positive number M'_n, there exists a code (ψ_n,ϕ_n) such that

$$\epsilon_n(\psi_n,\phi_n) \le \sum_x P^n(x) W_x^n\left\{\frac{W^n(y|x)}{W_{P^n}^n(y)} \le M'_n\right\} + \frac{|\psi_n|}{2M'_n}$$
$$|\psi_n| = M_n.$$

Lemma 2 (*Verdú-Han* [14]): For any channel $W^n(y|x)$ and any code (ψ_n,ϕ_n), there exists a probability distribution $P^n(x)$ satisfying the following condition. Any integer M'_n satisfies that

$$\epsilon_n(\psi_n,\phi_n) \ge \sum_x P^n(x) W_x^n\left\{\frac{W^n(y|x)}{W_{P^n}^n(y)} \le M'_n\right\} - \frac{M'_n}{|\psi_n|}. \tag{19}$$

Using Lemma 1, we show

$$\liminf_{n\to\infty} C_n(R_1 n|W^n) \le \lim_{\gamma\to+0} \inf_{\bar{P}} \underline{P}_I(R_1+\gamma|\bar{P},\bar{W}).$$

First, applying Lemma 1 to the case of $M_n = 2^{R_1 n}$ and $M'_n = 2^{(R_1+\gamma)n}$, we obtain

$$\epsilon_n(\psi_n,\phi_n) \le \sum_x P^n(x) W_x^n\left\{\log\frac{W^n(y|x)}{W_{P^n}^n(y)} \le (R_1+\gamma)n\right\} + \frac{1}{2\cdot 2^{\gamma n}},$$

which implies

$$\liminf_{n\to\infty} \epsilon_n(\phi_n,\psi_n) \le \underline{P}_I(R_1+\gamma|\bar{P},\bar{W}).$$

Hence, $\liminf_{n\to\infty} C_n(R_1 n|W^n) \le \underline{P}_I(R_1+\gamma|\bar{P},\bar{W})$. Taking the infimum $\inf_{\bar{P}}$ and the limit $\lim_{\gamma\to+0}$, we obtain the desired inequality.

Next, using Lemma 2, we prove $\lim_{\gamma\to+0}\inf_{\bar{P}}\underline{P}_I(R_1-\gamma|\bar{P},\bar{W}) \le \liminf_{n\to\infty} C_n(R_1 n|W^n)$. We apply Lemma 2 to a code (ψ_n,ϕ_n)

realizing the minimum average error probability $C_n(R_1 n|W^n)$. Then, we denote a sequence of distribution satisfying the condition in Lemma 2 by $\bar{P} = \{P^n(x)\}$. Choosing $M'_n = 2^{(R_1-\gamma)n}$, we obtain

$$\epsilon_n(\psi_n,\phi_n) \ge \sum_x P^n(x) W_x^n \left\{ \log \frac{W^n(y|x)}{W_{P^n}^n(y)} \le (R_1-\gamma)n \right\} - \frac{2^{(R_1-\gamma)n}}{|\psi_n|}.$$

Since $(2^{(R_1-\gamma)n})/|\psi_n| \to 0$, we have

$$\liminf_{n\to\infty} C_n(R_1 n|W^n) = \liminf_{n\to\infty} \epsilon_n(\psi_n,\phi_n) \ge \underline{P}_I(R_1-\gamma|\bar{P},\bar{W}),$$

which implies $\liminf_{n\to\infty} C_n(R_1 n|W^n) \ge \inf_{\bar{P}} \underline{P}_I(R_1-\gamma|\bar{P},\bar{W})$. Taking the limit $\lim_{\gamma\to+0}$, we obtain $\liminf_{n\to\infty} C_n(R_1 n|W^n) \ge \lim_{\gamma\to+0}\inf_{\bar{P}} \underline{P}_I(R_1-\gamma|\bar{P},\bar{W})$. Replacing $\liminf_{n\to\infty}$ by $\limsup_{n\to\infty}$, we can show (12). Then, we complete our proof.

The general formula for the second order asymptotics can be obtained in a similar way. For this purpose, we define

$$\begin{aligned}&\overline{P}_I(R_1,R_2|\bar{P},\bar{W})\\ &:= \limsup_{n\to\infty} \sum_{x\in\mathcal{X}_n} P^n(x) W_x^n \left\{ \log \frac{W^n(y|x)}{W_{P^n}^n(y)} \le R_1 n + R_2\sqrt{n} \right\}.\end{aligned}$$

Replacing $\limsup_{n\to\infty}$ by $\liminf_{n\to\infty}$, we can define another quantity $\underline{P}_I(R_1,R_2|\bar{P},\bar{W})$. Then, we obtain

$$\begin{aligned}\lim_{\gamma\to+0}\inf_{\bar{P}} \underline{P}_I(R_1,R_2-\gamma|\bar{P},\bar{W}) &\le \liminf_{n\to\infty} C_n(R_1 n + R_2\sqrt{n}|W^n)\\ &\le \lim_{\gamma\to+0}\inf_{\bar{P}} \underline{P}_I(R_1,R_2+\gamma|\bar{P},\bar{W}) \quad (20)\\ \lim_{\gamma\to+0}\inf_{\bar{P}} \overline{P}_I(R_1,R_2-\gamma|\bar{P},\bar{W}) &\le \limsup_{n\to\infty} C_n(R_1 n + R_2\sqrt{n}|W^n)\\ &\le \lim_{\gamma\to+0}\inf_{\bar{P}} \overline{P}_I(R_1,R_2+\gamma|\bar{P},\bar{W}). \quad (21)\end{aligned}$$

The proofs of (11) and (12) can be applied to the proofs of the above by replacing $M'_n = 2^{(R_1+\gamma)n}$ and $M'_n = 2^{(R_1-\gamma)n}$ by $M'_n = 2^{R_1 n+(R_2+\gamma)\sqrt{n}}$ and $M'_n = 2^{R_1 n+(R_2-\gamma)\sqrt{n}}$.

The second order asymptotics can be discussed in the same way as the first order asymptotics from the general viewpoint of information spectrum. However, even if the channel is stationary and memoryless, it is not so easy to calculate the quantity $\lim_{\gamma\to+0}\inf_{\bar{P}} \underline{P}_I(R_1,R_2-\gamma|\bar{P},\bar{W})$.

But, when the channel is additive, i.e., is given by (15) and $|\mathcal{X}_n| = d^n$, similar to the case of the first order asymptotics, we obtain

$$\inf_{\bar{P}} \overline{P}_I(R_1,R_2|\bar{P},\bar{W}) = 1 - \underline{P}_H(\log d - R_1, -R_2|\bar{Q}) \quad (22)$$

and

$$\inf_{\bar{P}} \underline{P}_I(R_1,R_2|\bar{P},\bar{W}) = 1 - \overline{P}_H(\log d - R_1, -R_2|\bar{Q}). \quad (23)$$

Hence, due to the central limit theorem, the independent and identical distribution $\bar{Q}$ of Q satisfies

$$\begin{aligned}\overline{P}_H(R_1,R_2|\bar{Q}) &= \underline{P}_H(R_1,R_2|\bar{Q})\\ &= \begin{cases} 1 & \text{if } R_1 > H(Q)\\ \Phi(R_2/\sqrt{V(Q)}) & \text{if } R_1 = H(Q)\\ 0 & \text{if } R_1 < H(Q), \end{cases} \quad (24)\end{aligned}$$

we obtain

$$\begin{aligned}&\lim_{n\to\infty} C_n(R_1 n + R_2\sqrt{n}|W^n)\\ &= \begin{cases} 0 & \text{if } R_1 < \log d - H(Q)\\ \Phi(R_2/\sqrt{V(Q)}) & \text{if } R_1 = \log d - H(Q)\\ 1 & \text{if } R_1 > \log d - H(Q) \end{cases} \quad (25)\end{aligned}$$

because $\Phi(R_2/\sqrt{V(Q)})\ 1 - \Phi(-R_2/\sqrt{V(Q)})$.

On the other hand, the derivation of (7) from (20) and (21) is more complicated when the channel is stationary and memoryless, but is non-additive.

In this case, we choose P^+ and P^- satisfying

$$I(P^+,W) = I(P^-,W) = C(W)$$

and

$$\begin{aligned}V^+(W) &= \sum_x P^+(x) \sum_y W_x(y) \left(\log \frac{W_x(y)}{W_P(y)} - D(W_x\|W_P) \right)^2\\ V^-(W) &= \sum_x P^-(x) \sum_y W_x(y) \left(\log \frac{W_x(y)}{W_P(y)} - D(W_x\|W_P) \right)^2.\end{aligned}$$

When $\bar{P}_+$ and $\bar{P}_-$ are the independent and identical distribution of P^+,

$$\begin{aligned}\overline{P}_I(R_1,R_2|\bar{P}_+,\bar{W}) &= \Phi(R_2/\sqrt{V^+(W)})\\ \overline{P}_I(R_1,R_2|\bar{P}_-,\bar{W}) &= \Phi(R_2/\sqrt{V^-(W)}).\end{aligned}$$

So, we obtain the inequality $\le$ in (7).

The opposite inequality can be shown by considering the average distribution concerning the permutation and employing the method of type [9]. A similar relation can be shown in the similar way in the additive Gaussian noise case and the energy constraint case [9].

The method in this section can be extended to the cases of source coding and uniform random number generation [10]. The second order asymptotic theory for channel coding has been studied first by Strassen [8] and recently done by Polyanskiy [15]. However, in order to treat the second order asymptotic theory in a unified viewpoint, the method by information spectrum is better than others.

V. Application to Quantum Key Distribution

The second order asymptotics is not only important for calculation of theoretical minimum error but also essential for security analysis of quantum key distribution [16], [17]. In the above discussion, we have explained the second order asymptotic theory for the minimum average error.

One might doubt whether the minimum average error is realizable. In fact, it is quite difficult to realize a code attaining the minimum average error. However, it is possible to realize the encoder

attaining the minimum average error if the decoder construction is not required. That is, if we accept the convenient assumption that we can ignore the decoding time, the second order asymptotic theory for channel coding has a realistic meaning. In fact, the evaluation of security for quantum key distribution is nothing than such a situation convenient for information theory.

In quantum key distribution, choosing the their basis randomly, the sender and the receiver can make the eavesdropper disable to access the basis information of the random bits. Hence, if there is no noise in the quantum communication, we can decide whether there exists the eavesdropper because the existence of the disagreement between the sender's and the receiver's bits implies the existence of eavesdropping. However, the real quantum communication has negligible noise.

Hence, we need to adapt the error correcting code for the basis to be used and the dual basis. In quantum key distribution, the leaking information can be evaluated only by the error probability of the dual basis. That is, if the error correction has been done for the dual basis perfectly, it is guaranteed that there is no leaking information.

If the receiver do not decode the received information on the used basis, the communication is not reliable. But, the dual basis does not require a decoding because only the the possibility of decoding is required for guaranteeing the security. Further, the real encoding for the dual basis requires very complicated quantum process, it can be replaced by application of hash function. That is, the computation cost for random coding for the dual basis is essentially realizable, and we do not have to care the computation cost for decoding.

Even in such a convenient setting, it is not allowable to choose n to be infinity. So, we have to treat the finite size case, e.g., $n = 10{,}000$. In such a situation, the second order asymptotic theory for channel coding is quite effective for the approximate evaluation of the error probability for the dual basis. Further, modifying the discussion on the second order asymptotics, we can derive a tighter upper bound for the average error probability on the dual basis, which is quite useful for quantitative evaluation of security of quantum key distribution [18].

VI. Application to Folklore in Source Coding

Many people have believed that, the compressed data obeys the uniform distribution when the original data sequence obeys subject to a biased distribution P and it is compressed up to the entropy rate $H(P)$. This is because the limit of compression rate equals the limit of the generation rate of the uniform random number. However, this argument had not been mathematically formulated nor discussed sufficiently for a long time. The first analysis has been done by Han [19]. After his analysis, Gray [20] has done a different type analysis for this topic based on a different criterion.

Before we treat the relation with the second order asymptotics, we summarize the result for the second order asymptotics with the fixed-length date compression and the uniform random number generation. In the following, we focus on the sequence of n random numbers $\mathbf{X}^n := (X_1, \ldots, X_n)$ that independently obey the distribution P^X on $\mathcal{X}$. Then, we define the encoder as a map ϕ_n from $\mathcal{X}^n$ to the set $\mathcal{M}_n := \{1, \ldots, M_n\}$, and the decoder as a map ψ_n from $\mathcal{M}_n$ to $\mathcal{X}^n$. The average error probability for the data compression is given by $\epsilon_n(\phi_n, \psi_n) := \sum_{x \in \mathcal{X}^n : \psi_n \circ \phi_n(x) \neq x} (P^X)^n(x)$. We measure the quality as the uniform random number by the half of the variational distance[1] $d(P^{\phi_n(X^n)}, P^{\phi_n, U})$ between the distribution $P^{\phi_n(X^n)}$ of the generated random number and the uniform distribution $P^{\phi_n, U}$ on the image of ϕ_n. Then, we define the minimum average error probability with a fixed compression length L as

$$R_n(L|P^X) := \min_{(\phi_n, \psi_n)} \{\epsilon_n(\phi_n, \psi_n) | \log |\phi_n| \le L\}$$

and the variational distance with a fixed generation length L as

$$S_n(L|P^X) := \min_{\phi_n} \left\{ d_1(P^{\phi_n(X^{(n)})}, P^{\phi_n, U})/2 \,|\log|\phi_n| \ge L \right\}.$$

The limits with the second order asymptotics are given as

$$\lim_{n\to\infty} R_n(R_1 n + R_2\sqrt{n}\,|P^X) = \begin{cases} 1 & \text{if } R_1 < H(P^X) \\ 1 - \Phi(R_2/\sqrt{V(P^X)}) & \text{if } R_1 = H(P^X) \\ 0 & \text{if } R_1 > H(P^X) \end{cases} \quad (26)$$

$$\lim_{n\to\infty} S_n(R_1 n + R_2\sqrt{n}\,|P^X) = \begin{cases} 0 & \text{if } R_1 < H(P^X) \\ \Phi(R_2/\sqrt{V(P^X)}) & \text{if } R_1 = H(P^X) \\ 1 & \text{if } R_1 > H(P^X) \end{cases} \quad (27)$$

These formulas can be shown from the same framework as the channel coding via the method of information spectrum [10].

Next, we introduce Han [19]'s approach to folklore insource coding. He formulated this problem based on normalized value of the KL-divergence between the uniform distribution and the distribution of the generated random number. That is, for the above code (ϕ_n, ψ_n), he focused on the value $(1/n)D(P^{\phi_n(X^n)} \| P^{\phi_n, U})$. He proved the following theorem concerning the fixed-length data compression.

Theorem 3 (Han [19]): When a fixed data compression (ϕ_n, ψ_n) satisfies that $\epsilon_n(\phi_n, \psi_n)$ goes to zero and $(1/n)\log M_n$ goes to $H(P^X)$, we obtain

$$\frac{1}{n} D(P^{\phi_n(X^n)} \| P^{\phi_n, U}) \to 0. \quad (28)$$

This theorem can be easily shown by noticing

$$D(P^{\phi_n(X^n)} \| P^{\phi_n, U}) = H(P^{\phi_n, U}) - H(P^{\phi_n(X^n)}) = \log M_n - H(P^{\phi_n(X^n)}). \quad (29)$$

Since $\epsilon_n(\phi_n, \psi_n)$ equals the half of the variational distance between $P^{\psi_n \circ \phi_n(X^n)}$ and P^{X^n}, Fannes' inequality yields that

$$|H(P^{X^n}) - H(P^{\psi_n \circ \phi_n(X^n)})| \le 2\epsilon_n(\phi_n, \psi_n)\log|\mathcal{X}|^n, \quad (30)$$

which implies $(1/n)HP^{\psi_n \circ \phi_n(\mathbf{X}^n)} \to H(P^{\mathbf{X}})$. Since the maps ψ_n and ϕ_n decreases the entropy we obtain $H(P^{\psi_n \circ \phi_n(\mathbf{X}^n)}) \le H(P^{\phi_n(X^n)}) \le H(P^{X^n}) = nH(P^{\mathbf{X}})$. Thus, we can confirm that $(1/n)H(P^{\phi_n(X^n)}) \to H(P^{\mathbf{X}})$. Thus, (29) implies (28).

[1] $d_1(P, Q) := \Sigma_x |P(x) - Q(x)|$.

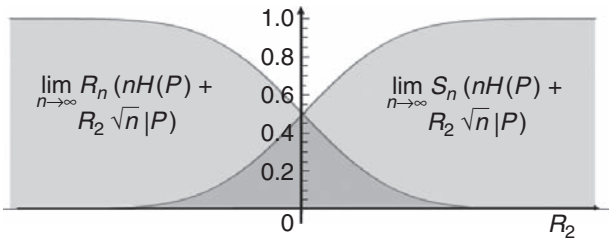


Fig. 3. The graph of $\lim_{n\to\infty} R_n(H(P^X)n + R_2\sqrt{n}|P^X)$ and $\lim_{n\to\infty} S_n(H(P^X)n + R_2\sqrt{n}|P^X)$.

We have easily shown the above theorem, however, it is not so clear whether the normalized KL-divergence should be used as the criteria for the uniform random number.

If the variation distance between the uniform distribution and the distribution of the generated random number does not go to zero, we can distinguish the generated random number from the uniform random number. So, in order to claim that the generated random number is close to the uniform random number, we should show that the above variation distance goes to zero. Alternatively, we should show that the KL-divergence goes to zero because the convergence concerning the KL-divergence is stronger than the convergence concerning the variational distance due to Pinsker's inequality [21].

If we focus on the first order asymptotics, we cannot deny the existence of a code such that the variational distance from the uniform random number goes to zero and the average error probability for the data compression goes to zero. This is because there exists a common possible rate $H(P^X)$ for both conditions. However, the second order asymptotics denies the existence of such a code because there exist no rate for the second order asymptotics satisfying both conditions due to (26) and (27). That is, in order that the variational distance goes to zero, the second rate R_2 must be $-\infty$, and in order that the average error for data compression goes to zero, the second rate R_2 must be $+\infty$. As is illustrated in Fig. 3, It is impossible to satisfy the both conditions. Therefore, when the data can be correctly recovered in the data compression, the compressed data cannot go to the uniform distribution in the sense of the variational distance [10][2]. The above fact means that the second order analysis discovers the detail behavior behind of the first order asymptotics.

VII. Application to Security

The second order asymptotics reveals the problem for traditional information-theoretical security analysis based on the first order asymptotics. The previous section discusses the relation between the compressed random number and the uniform random number. This topic relates to the generation of the secret random number as well as the uniform random number. Now, we consider how to distill secret random number from a random number A leaked to the eavesdropper as a partially correlated random number E.

[2] This fact has been shown independently by Han [19] during the review process of [10]. The paper [10] has been written based on the stimulation by the conference version of Han [19], which does not treat the variational distance.

In this case, applying Hash function to the random number A, we can distill a random number B that is almost independent of the other random number E. This process is called privacy amplification, and has been studied from the community cryptography theory. On the other hand, due to Slepian-Wolf theorem [22], if the additional side information E is available, the random number A can be compressed up to the conditional entropy rate $H(A|E)$. Let B be the random number obtained by the compression up to the conditional entropy rate $H(A|E)$ based on the leaking information E Then, if the folklore in source coding is valid and the random number A can be decoded from B with E, B looks the uniform distribution for respective values of E. That is, the eavesdropper cannot obtain any information concerning B from E.

However, the folklore in source coding is valid only under the normalized KL-divergence criterion. It is not valid under the non-normalized KL-divergence criterion. This fact can be extended to the case when the additional information E exists [23]. Therefore, the compressed random number B has no correlation with E under the normalized mutual information criterion. However, we cannot say that it has no correlation with E under the non-normalized mutual information criterion.

Employing the separation coding by Slepian-Wolf [22], Ahlswede, Csiszar [24] and Muramatsu [25] treated the secret key distillation from two correlated random variables A and A' that are partially eavesdropped as another random variable E. For a simplicity, in the following, we consider the case $A = A'$, i.e., the case when the error correction is not required. Under this limited case, their method proposed is simplified as follows. We convert the initial random number A to the pair of the final random number B and the dummy random number C. In their method, we keep B as the final random number.

Assume that the random number A can be perfectly recovered from the pair of B and E and $H(C) = H(C|B) = I(A:E)$, which is equivalent with the condition that the random number C can be perfectly recovered from the random number E and $H(C) = H(C|B) = I(A:E)$. Then, we obtain $I(A:E|B) = H(A|B) + H(E|B) - H(AE|B) = H(BC|B) + H(E|B) - H(BEE|B) = H(C|B) + H(E|B) - H(E|B) = H(C|B)$, which implies that $I(B:E) = I(A:E) - I(A:E|B) = 0$, i.e., B is perfectly independent of E. However, in our situation, it is impossible to satisfy the condition $H(C) = I(A:E)$ perfectly. It is possible to set the conditional entropy rate $H(C|B)$ and the average decoding error probability for C from E to $I(A:E)$ and zero, asymptotically. In this case, we can show that the normalized mutual information between B and E goes to zero asymptotically. Alternatively, using the second order asymptotics, Watanabe et al [23] showed that the mutual information between B and E does not converge to zero but behaves with order $\sqrt{n}$ when the dummy random number C can be decoded. That is, the strong security does not hold with the above construction.

Therefore, the security analysis based on the first order asymptotics is effective only under the normalized mutual information, but the higher order asymptotics is required for a smaller order analysis of the mutual information. As is shown by Watanabe et al [23], the mutual information behaves with the order $\sqrt{n}$ under the above construction. Any higher analysis than Watanabe et al [23] does not improve the order analysis of the mutual information.

Thus, such a higher analysis is not required for analysis of the convergence of the mutual information.

Using the second order analysis, we can show that the method via error correction for C cannot generate the secure random number under the non-normalized mutual information. Hence, we need to directly evaluate the variational distance from the uniform distribution or the non-normalized mutual information without the error correction for C.

While the method via the error correction for C is familiar to researchers in information theory, the direct evaluation for the variational distance or the non-normalized mutual information requires a completely new method. As for a method directly evaluating these values, the privacy amplification theorem has been studied from the community of cryptography theory. Initially, this theorem has been established with the Renyi entropy with order 2 [26], [27]. Renner and his collaborators extended it to the version with the smooth min entropy [29], [28]. Recently, the author extended it to the version with Renyi entropy with order $1+s$, which directly yields the exponential decreasing rate of the mutual information when the generation rate is less than the conditional entropy rate [30], [31].

The same problem happens for the security analysis in wire-tap channel. When the normalized mutual information is adopted as the security criterion, the security can be shown via the error correction for the dummy random variable C. However, it is impossible to show the the security under the non-normalized mutual information criterion, i.e., the strong security based on this method because the mutual information behaves as the order $\sqrt{n}$ [32]. As such a method guaranteeing the strong security, the author and the collaborator proposed an application of privacy amplification theorem or channel resolvability [33] to wire-tap channel [30], [31], [34], [35], [36], [37], [38], [39], [40][3]. Therefore, the second order asymptotics enables us to evaluate the detail order analysis that cannot be treated by the traditional the first order asymptotics. That is, the second order asymptotics reveals a kind of security hole behind of the traditional approach. We can expect further reveals of the problem caused by the first order asymptotics.

IX. Conclusions

We have explained the second order asymptotics theory for the channel coding, fixed-length data compression, and the uniform random number generation. They can be treated from a unified viewpoint by employing the method of information spectrum due to the generality of information spectrum.

The method of information spectrum can be applied not only to the second order asymptotics but also to other topics. For example, it is known that in the case of quantum communication with coherent states, the number of transmittable bits increases up to infinity even if the total energy is fixed when the number of used modes increases up to infinity. Such a phenomena does not happen in the classical case. The method of information spectrum enables us to resolve the asymptotic analysis for such a case [46]. Hence, we can expect a further and wider variety of applications of the method of information spectrum.

[3]The strong security based on channel resolvability has been given by Csiszar [42] firstly, and applied to the quantum case by Devetak [43],Winter et al [44], [45], independently. However, these researches have not been paid attention to in the classical information community. Then, the method based on the error correction of the dummy has been mainly used for the security analysis of the classical wire-tap channel. The main reason seems that the initial study by Wyner [41] employs this method. Note that the method via error correction for phase error is completely different from this method. The method via error correction for phase error can guarantee the strong security because the mutual information can be directly evaluated via the phase error [18], [16].

X. Acknowledgments

The author is grateful to Professors Te Sun Han, Hiroshi Nagaoka for explaining information spectrum. He is also grateful to Professors Tomohiko Uyematsu and Ryutaroh Matsumoto for informing the reference [8] and the references [23], [32], [20], respectively. He is also grateful to anonymous referees and editors of the papers [9], [10] for helpful comments. He also thanks Dr. Ke Li, Dr. Akihito Soeda, and Mr. Wataru Kumagai for helpful comments. This study was partially supported by MEXT through a Grant-in-Aid for Scientific Research on Priority Area "Deepening and Expansion of Statistical Mechanical Informatics (DEX-SMI)", No. 18079014 a MEXT Grant-in-Aid for Young Scientists (A) No. 20686026, and a MEXT Grant-in-Aid for Scientific Research (A) No. 23246071. The Centre for Quantum Technologies is funded by the Singapore Ministry of Education and the National Research Foundation as part of the Research Centres of Excellence programme.

Author

M. Hayashi is with Graduate School of Information Sciences, Tohoku University, Aoba-ku, Sendai, 980-8579, Japan and Centre for Quantum Technologies, National University of Singapore, 3 Science Drive 2, Singapore 117542. (e-mail: hayashi@math.is.tohoku.ac.jp)

References

[1] B. Efron. Defining the curvature of a statistical problem (with applications to second order efficiency). *Ann. Statist.*, 3(6):1189–1242, 1975.

[2] S. Amari and H. Nagaoka, *Methods of information geometry,* Translated from the 1993 Japanese original by Daishi Harada. Translations of Mathematical Monographs, 191. American Mathematical Society, Providence, RI; Oxford University Press, Oxford, 2000.

[3] I. Kontoyiannis, "Second-order noiseless source coding theorems," *IEEE Trans. Inform. Theory*, 43, 1339–1341 (1997); I. Kontoyiannis, "Pointwise redundancy in lossy data compression and universal lossy data compression." *IEEE Trans. Inform. Theory*, 46, 136–152, (2000).

[4] B. S. Clarke and A. R. Barron, "Jeffreys' prior is asymptotically least favorable under entropy risk," *Journal of Statistical Planning and Inference*, 41, 37–61 (1994).

[5] P. Flajolet and W. Szpankowski, "Analytic variations on redundancy rates of renewal processes," *IEEE Trans. Inform. Theory*, 48, 2911–2921 (2002).

[6] M. Drmota and W. Szpankowski, "Precise minimax redundancy and regret," *IEEE Trans. Inform. Theory*, 50, 2686–2707 (2004).

[7] E. Figueroa and C. Houdre, "On the Asymptotic Redundancy of Lossless Block Coding With Two Codeword Lengths," *IEEE Trans. Inform. Theory*, 51, 688–692 (2005).

[8] V. Strassen, "Asymptotische Abschätzugen in Shannon's Informations theorie," In Transactions of the Third Prague Conference on Information Theory etc, 1962. Czechoslovak Academy of Sciences, Prague, pp. 689–723.

[9] M. Hayashi, "Information Spectrum Approach to Second-Order Coding Rate in Channel Coding," IEEE Transactions on Information Theory, Vol. 55, No. 11, 4947–4966 (2009);
[10] M. Hayashi, "Second-Order Asymptotics in Fixed-Length Source Coding and Intrinsic Randomness," IEEE Transactions on Information Theory, Vol. 54, 4619–4637 (2008);
[11] T.-S. Han, *Information-Spectrum Methods in Information Theory*, (Springer, Berlin, 2003). (Originally published by Baifukan 1998 in Japanese)
[12] R. Nomura, T.-S. Han, "Second-Order Resolvability, Intrinsic Randomness, and Fixed-Length Source Coding for Mixed Sources,"ar Xiv:1106.1879.
[13] R. G. Gallager, *Information Theory and Reliable Communication*, John Wiley & Sons, 1968.
[14] S. Verdú and T.-S. Han, "A general formula for channel capacity," *IEEE Trans. Inform. Theory*, vol. 40, 1147–1157, 1994.
[15] Y. Polyanskiy, H.V. Poor, S. Verdú, "Channel coding rate in the finite blocklength regime," IEEE Trans. Inform. Theory, Vol. 56, 2307–2359 (2010).
[16] M. Hayashi, "Practical Evaluation of Security for Quantum Key Distribution," *Phys. Rev. A*, 74, 022307 (2006).
[17] M. Hayashi, "Upper bounds of eavesdropper's performances in finite-length code with the decoy method," Physical Review A, Vol. 76, 012329 (2007); Physical Review A, Vol. 79, 019901(E) (2009)
[18] M. Hayashi, T. Tsurumaru "Simple and Tight Security Analysis of the Bennett-Brassard 1984 Protocol with Finite Key Lengths," arXiv:1107.0589.
[19] T.-S. Han, "Folklore in source coding: information-spectrum approach," *IEEE Transactions on Information Theory*, Vol. 51, 747–753, (2005).
[20] R. M. Gray, "Source Coding and Simulation," IEEE Information Theory Society Newsletter Vol. 58, No. 4, p.1, 5–11 (2008).
[21] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Budapest, Hungary: Akadémiai Kiadó, 1981.
[22] D. Slepian and J.K. Wolf, "Noiseless coding of correlated information sources," IEEE Trans. Inform. Theory, vol. 19, no. 4, pp. 471–480, July 1973.
[23] S. Watanabe, R. Matsumoto, and T. Uyematsu "Strongly Secure Privacy Amplification Cannot Be Obtained by Encoder of Slepian-Wolf Code," IEICE Trans. Fundamentals, A-93, 1650–1659 (2010).
[24] R. Ahlswede and I. Csisz'ar, "Common randomness in information theory and cryptography part I: Secret sharing," IEEE Trans. Inform. Theory, vol. 39, no. 4, pp. 1121–1132, July 1993.
[25] J. Muramatsu, "Secret key agreement from correlated source outputs using low density parity check matrices," IEICE Trans. Fundamentals, vol. E89-A, no. 7, pp. 2036–2046, 2006.
[26] J. Håstad, R. Impagliazzo, L. A. Levin, and M. Luby, "A Pseudorandom Generator from any One-way Function," SIAM J. Comput. 28, 1364 (1999)
[27] C.H.Bennett, G. Brassard, C. Crepeau, and U.M. Maurer, "Generalized privacy amplification," *IEEE Trans. Inform. Theory*, vol. 41, 1915–1923, 1995.
[28] R. Renner and S. Wolf, "Simple and tight bounds for information reconciliation and privacy amplification," in Proc. ASIACRYPT, 2005, vol. 3788, pp. 199?216, Lecture Notes in Computer Science, Springer-Verlag.
[29] R. Renner, "Security of Quantum Key Distribution," PhD thesis, Dipl. Phys. ETH, Switzerland, 2005. arXiv:quantph/0512258.
[30] M. Hayashi, "Exponential decreasing rate of leaked information in universal random privacy amplification," IEEE Transactions on Information Theory, Vol. 57, No. 6, 3989–4001, (2011).
[31] M. Hayashi "Tight exponential evaluation for information theoretical secrecy based on universal composablity," arXiv:1010.1358 (2010).
[32] H. Mahdavifar, A. Vardy, "Achieving the Secrecy Capacity of Wiretap Channels Using Polar Codes," IEEE Transactions on Information Theory, 57, 6428–6443 (2011).
[33] T. S. Han and S. Verdú, "Approximation theory of output statistics," IEEE Trans. Inf. Theory, vol. 39, no. 3, pp. 752–772, May 1993.
[34] M. Hayashi, "General non-asymptotic and asymptotic formulas in channel resolvability and identification capacity and its application to wire-tap channel," IEEE Transactions on Information Theory, Vol. 52, No. 4, 1562–1575 (2006).
[35] M. Hayashi, R. Matsumoto, "Universally Attainable Error and Information Exponents, and Equivocation Rate for the Broadcast Channels with Confidential Messages,"
[36] R. Matsumoto, M. Hayashi, "Secure Multiplex Network Coding," Proceedings of 2011 International Symposium on Network Coding (NetCod), Beijing, China, 25–27 July 2011.
[37] R. Matsumoto, M. Hayashi, "Secure Multiplex Coding with a Common Message," Proceedings of 2011 IEEE International Symposium on Information Theory (ISIT 2011), Saint-Petersburg, Russia, July, 31–August, 5, 2011, pp 1965–1969.
[38] R. Matsumoto, M. Hayashi, "Strong security and separated code constructions for the broadcast channels with confidential messages," arXiv:1010.0743 (2010).
[39] M. Hayashi, R. Matsumoto, "Construction of wiretap codes from ordinary channel codes," Proc. 2010 IEEE ISIT, pp. 2538–2542, Austin, Texas, USA, June 13–18, 2010.
[40] M. Hayashi, Quantum Information: An Introduction, Springer (2006).
[41] A. D. Wyner, "The wire-tap channel," *Bell. Sys. Tech. Jour.*, vol. 54, 1355–1387, 1975.
[42] I. Csiszar, "Almost independence and secrecy capacity," *Problems of Information Transmission*, vol. 32, no. 1, pp. 40–47, 1996.
[43] I. Devetak, "The private classical information capacity and quantum information capacity of a quantum channel," *IEEE Trans. Inform. Theory*, vol. 51(1), 44–55, 2005.
[44] Andreas Winter, Anderson C. A. Nascimento and Hideki Imai, "Commitment Capacity of Discrete Memoryless Channels," Proceedings of Cryptography and Coding 9th IMA International Conference, Cirencester, UK, December 16–18, 2003. Proceedings, Lecture Notes in Computer Science, 2003, Volume 2898/2003, 35–51.
[45] N. Cai, A. Winter and R. W. Yeung, "Quantum privacy and quantum wiretap channels," *Problems of Information Transmission*, Volume 40, Number 4, 318–336, (2004)
[46] M. Hayashi, "Capacity with energy constraint in coherent state channel," *IEEE Transactions on Information Theory*, Vol. 56, No. 8, 4071–4079, (2010).

Ants and Bits

Plenary talk presented at the 2011 IEEE International Symposium of Information Theory, St. Petersburg, Russia.

Zhanna Reznikova and Boris Ryabko

Abstract. Ants have always been helping people to solve various problems. Everybody remembers how they sorted seeds for Cinderella. For the IT community, ants have helped to show that Information Theory is not only an excellent mathematical theory but that many of its results can be considered laws of Nature. Reciprocally, we helped ants to be distinguished among other "intellectuals" such as counting primates, crows and parrots as one of the smartest species [1, 2]. Our long-term experimental study on ant "language" and intelligence were fully based on fundamental ideas of Information Theory, such as the Shannon entropy, the Kolmogorov complexity, and the Shannon's equation connecting the length of a message l and its frequency of occurrence p, i.e., $l = -\log p$. This approach enabled us to discover a developed symbolic "language" in highly social ant species based on their ability to transfer the abstract information about remote events and to estimate the rate of information transmission. We also succeeded to reveal important properties of ants' intelligence. These insects appeared to be able to grasp regularities and to use them for "compression" of data they communicate to each other. They can also transfer to each other the information about the number of objects and can even add and subtract small numbers in order to optimize their messages.

Introduction

From time immemorial, people have been dreaming about understanding animal "languages"- a dream with which many legends are associated. The title of the book of the famous ethologist Konrad Lorenz, King Solomon's Ring (1952), refers to the legend about King Solomon who possessed a magical ring that gave him the power of speaking with animals. However, decoding the function and meaning of animal communications is a notoriously difficult problem. A bottleneck here is the low repeatability of standard living situations, which could give keys for cracking animals' species-specific codes. Up to now, there are only two types of natural communication systems that have been partly deciphered: the fragments of honeybees' "dance language", and acoustic signalization in vervet monkeys and several other species (see [3] for a review). In both types of communications, expressive and distinctive signals correspond to repeatable and frequently occurring situations in the context of animals' life. The problem of cracking animals' codes have become especially attractive since the great "linguistic" potential was discovered in several highly social and intelligent species by means of intermediary artificial languages. Being applied to apes, dolphins and gray parrots, this method has revealed astonishing mental skills in the subjects [4, 5, 6]. However, surprisingly little is known yet about natural communication systems of those species that were involved in language-training experiments based on adopted human languages. Explorers of animal "languages" thus have met a complex problem of resolving the contradiction between their knowledge about significant "linguistic" and cognitive potential in some species and limitations in understanding their natural communications.

We have suggested to apply ideas of Information Theory for studying natural communications of animals, that is, not to decipher signals, but to investigate the very process of information transmission by measuring time duration which the animals spend on transmitting messages of different lengths and complexities.

Ants of highly social species are good candidates for studying general rules of cognitive communication. There are more than 12000 ant species on Earth, and the great majority of them use relatively simple forms of communication such as odor trails, tandem running, and so on. Only a few highly social species belong to the elite club of rare "cognitive specialists", and among them are several species of red wood ants (*Formica rufa* group), with their big anthills "boiling" with hundreds of thousands of active individuals.

To reveal the power of ants' "language" we used two main notions of Information Theory, that is, (1) the quantity of information, and (2) the duration of time spent by the agents for transmitting 1 bit. This approach based on the "binary tree" experimental paradigm [7] enabled us to estimate the rate of information transmission in ants and to reveal that these insects are able to grasp regularities and to use them to compress information. The other series of experiments was based on the Shannon's equation connecting the length of a message (l) and its frequency (p), i.e., $l = -\log p$, for rational communication systems. Applying this concept, we demonstrated that ants are able to transfer to each other the information about the number of objects, and they even can add and subtract small numbers in order to optimize their messages [2].

The first results of this long-term work were reported at the International Information Theory Symposium 1984 in Tashkent *(the binary tree experiments) and at the ISIT-1994 in Norway (the ability to add and subtract small numbers), and then the obtained data were discussed at many international conferences and published in biological and mathematical journals (see, for example, [1, 4, 5]).

How to ask ants to transmit some bits of information to each other

The experimental paradigm of our approach is simple. All we need to do is to establish a situation where ants must transfer a specific amount of information to each other. The crucial idea of the first scheme of experiments is that we know exactly the quantity of information to be transferred and the time needed to do it. To organize the process of information transmission between ants, a special maze has been used, called a "binary tree" , where the number and sequence of turns towards the goal correspond to the amount of information to be transferred (Fig. 1).

In the laboratory we used fragments of ant colonies of about 2000 specimens each, and all active ants were labeled with color marks. Ants were housed in transparent artificial nests, so that their movements and contacts were observable. The laboratory colonies were

*A well-known Russian Information Theorist Yuri L. Sagalovitch related that year, he bought one of famous Tashkent big melons, and decided to have a rest on a bench, looking through the Symposium Proceedings. While reading about ants, he got carried away, and his melon was stolen.

found to include teams of constant membership which consisted of one scout and three to eight recruits (foragers): the scout mobilized only members of its team to the food. The composition of the teams was revealed during special run-up experiments. During the main course of experiments, in each trial one of the scouts was placed on a certain leaf of the binary tree that contained a trough with the food, and then it returned to the nest by itself. Returning to the group of foragers, the scout contacted one to four foragers in turn (Fig. 2). The duration of the contacts was measured every time.

All experiments were so devised as to eliminate all possible cues that could help the ants to find the food, except their information contact with the scout. To avoid the use of an odor track, the experimental set-up was replaced by an identical one when the scout was in the nest or on the arena contacting its group (see Fig. 3). All troughs in the fresh maze contained only water to avoid the possible influence of the smell of syrup. If the group reached the correct point, they were immediately presented with the food. The scout had to make up to four trips before it was able to mobilize its group of foragers. After the scout had contacted its team, it was isolated in a separate container for a while, and the foragers had to search for the food by themselves.

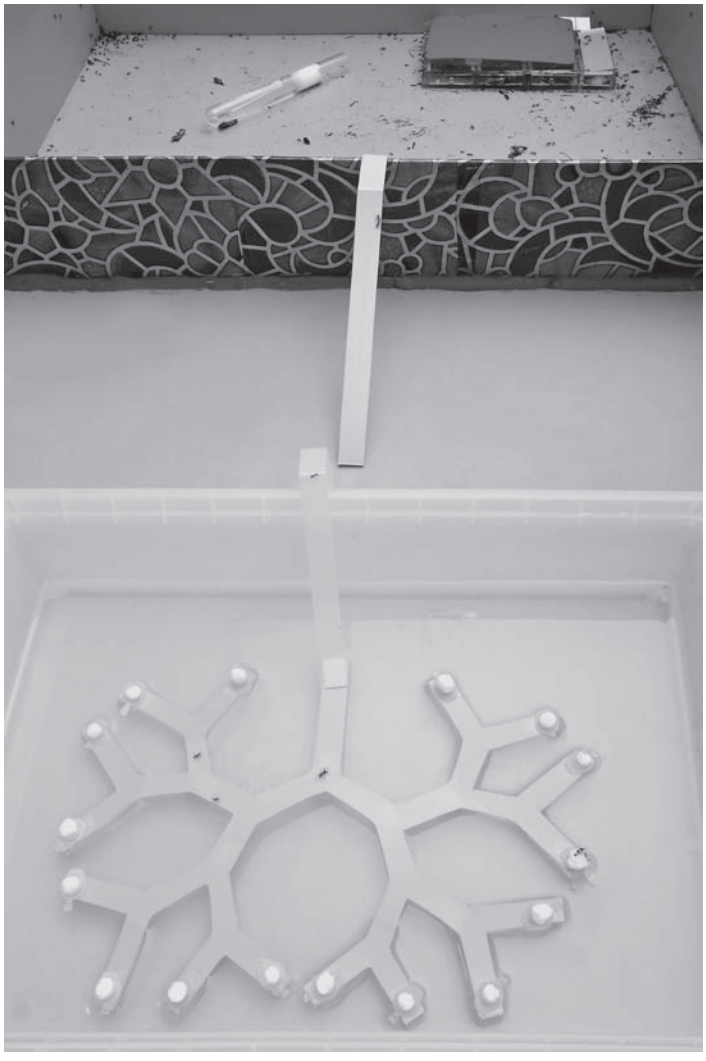


Fig. 1. A laboratory arena deviled into two parts, containing an artificial ant nest and a binary tree maze placed in a bath with water. Photo by Nail Bikbaev.

The experiments based on Shannon entropy present a situation in which, in order to obtain food, the ants have to transmit certain information which is quantitatively known to the researcher. This information concerns the sequence of turns towards a trough with syrup. The laboratory maze "binary tree" is used where each "leaf" of the tree ends with an empty trough with the exception of one filled with syrup. The leaf on which to place the filled trough was chosen randomly by tossing a coin for each fork in the path. The simplest design is a tree with one fork and two leaves, that is, a Y-shaped maze. It represents one binary choice which corresponds to one bit of information. In this situation a scouting animal should transmit one bit of information to other individuals: to go to the right (R) or to the left (L). In other experiments the number of forks of the binary tree increased to six. Hence, the number of bits necessary to choose the correct way is equal to the number of forks, that is, turns to be taken (Figure 1 shows a labyrinth with 3 forks). In total, 335 scouts along with their teams were used in all experiments with the binary tree, and each scout took part in tens of trials.

The binary tree and ants' language

Before analyzing ants' "linguistic potential" we considered the evidence of information transmission from the scouts to the foragers. The statistical analysis of the number of faultless findings of the goal was carried out by comparing the hypothesis H_0 (ants find the leaf containing the food by chance) with the hypothesis H_1 (they find the goal thanks to the information obtained), proceeding from the fact that the probability of finding the correct way by chance when the number of forks is i is $(1/2)^i$. We analyzed different series of experiments (338 trials in sum), separately for 2, 3, 4, 5, and 6 forks. In all cases H_0 was rejected in favor of H_1; $P < 0.001$, thus unambiguously demonstrating information transmission from scouts to foragers (see details in [1]).

In order to evaluate the rate of information transmission in ants, let us note that the quantity of information (in bits) necessary to choose the correct route in the maze equals i, the depth of the tree (the number of turns to be taken), that is, $\log 2\ n$ where n is the number of leaves. The obtained results have shown that the duration of the contacts between the scouts and foragers (t) is $ai + b$, where i is the number of turns (the depth of the tree), a is the time duration required for transmitting one bit of information, and b is an introduced constant, reflecting the fact that ants can transmit information not related directly to the task, for example, the simple signal "food". Besides, it is not ruled out that a scout ant transmits, in some way, the information on its route to the nest,

Fig. 2. A scouting ant contacting with members of its team. Photo by Nail Bikbaev.

using acoustic or some other means of communication. The rate of information transmission (a) derived from the equation $t = ai + b$ was about 1 minute per bit in three ant species, which is at least 10 times smaller than in humans.

Another series of experiments with the binary tree was inspired by the concept of Kolmogorov complexity and was designed to check whether ants possess such an important property of intelligent communications as the ability to grasp regularities and to use them for encoding and "compressing" information. This concept is applied to words (or text) composed of the letters of any alphabet, for example, of an alphabet consisting of two letters: L and R. We interpret a word as a sequence of left (L) and right (R) turns in a maze. Informally, the Kolmogorov complexity of a word (and its uncertainty) equates to its most concise description. For example, the word "LLLLLLLL" can be represented as "8 L", the word "LRLRLRLR" as "4LR", while the "random" word of shorter length "LRRLRL" probably cannot be expressed more concisely, and this is the most complex of the three.

We analyzed the question of whether ants can use simple regularities of a "word" to compress it. It is known that Kolmogorov complexity is not algorithmically computable. Therefore, strictly speaking, we can only check whether ants have a "notion" of simple and complex sequences. In our binary tree maze, in human perception, different routes have different complexities. We applied a statistical test in order to examine whether the time for transmission of information by ants depends on its complexity. We considered two hypotheses. The main hypothesis is H_0, that is, the time for transmission of information does not depend on the complexity of the "text". The alternative hypothesis is H1 that this time actually depends on the complexity of the "text". The hypothesis H_0 was rejected ($P = 0.01$), thus showing that the more time ants spent on the information transmission, the more complex - in the sense of Kolmogorov complexity - was the message (see details in [1]). It is worth to note that ants began using regularities to compress only quite large "texts". They spent from 120 to 220 sec. to transmit information about random turn patterns on the maze with 5 and 6 forks and from 78 to 135 sec. when turn patterns were regular. There was no essential difference when the length of sequences was less than 4.

Ideas of information theory and numerical competence in ants

Numerical competence is one of the main intriguing domains of animal intelligence. Recent studies have demonstrated some species as being able to judge about numbers of stimuli, including things, and sounds, and maybe smells (see [2] for a review). For example, lions can count roaring that comes from individuals who are not members of the pride; honey bees are able to use the number of landmarks as one of the criteria in searching for food sources. There are many other examples that come from different animal species, from mealy beetles to elephants; however, we are still lacking an adequate "language" for comparative analysis. The main difficulty in comparing numerical abilities in humans and other species is that our numerical competence is closely connected with abilities for language usage and for symbolic representation.

We suggested a new experimental paradigm which is based on ideas of information theory and is the first one to exploit natural communicative systems of animals. In our experiments ant scouts were required to transfer to foragers in a laboratory nest the information about which branch of a special "counting maze" they had to go to in order to obtain syrup. "Counting maze" is a collective name for several variants of set-ups (Fig. 4). The experiments were based on a procedure similar to the binary tree study. The main idea of this experimental paradigm is that experimenters can judge how ants represent numbers by estimating how much time individual ants spend on "pronouncing" numbers, that is, on transferring information about index numbers of branches.

Fig. 3. Dr. Natalia Azarkina (our guest from Moscow University) is marking ants by paint. One can see several extra mazes near the arena. Photo by Zhanna Reznikova.

The findings concerning number-related skills in ants are based on comparisons of duration of information contacts between scouts and foragers which preceded successful trips by the foraging teams. In total, 32 scout-foragers teams worked in three kinds of set-ups. It turned out that the relation between the index number of the branch (j) and the duration of the contact between the scout and the foragers (t) is well described by the equation $t = c\ j + d$ for different set-ups which are characterized by different shapes, distances between the branches and lengths of the branches. The values of parameters c and d are close and do not depend either on the lengths of the branches or on other parameters.

It is interesting that quantitative characteristics of the ants' "number system" seem to be close, at least outwardly, to some archaic human languages: the length of the code of a given number is proportional to its value. For example, the word "finger" corresponds to 1, "finger, finger" to the number 2, "finger, finger, finger" to the number 3 and so on. In modern human languages the length of the code word of a number j is approximately proportional to $\log j$ (for large j's), and the modern numeration system is the result of a long and complicated development.

An experimental scheme for studying ants' "arithmetic" skills based on a fundamental idea of information theory, which is that in a "reasonable" communication system the frequency of usage of a message and its length must correlate. The informal pattern is quite simple: the more frequently a message is used in a language, the shorter is the word or the phrase coding it. This phenomenon is manifested in all known human languages.

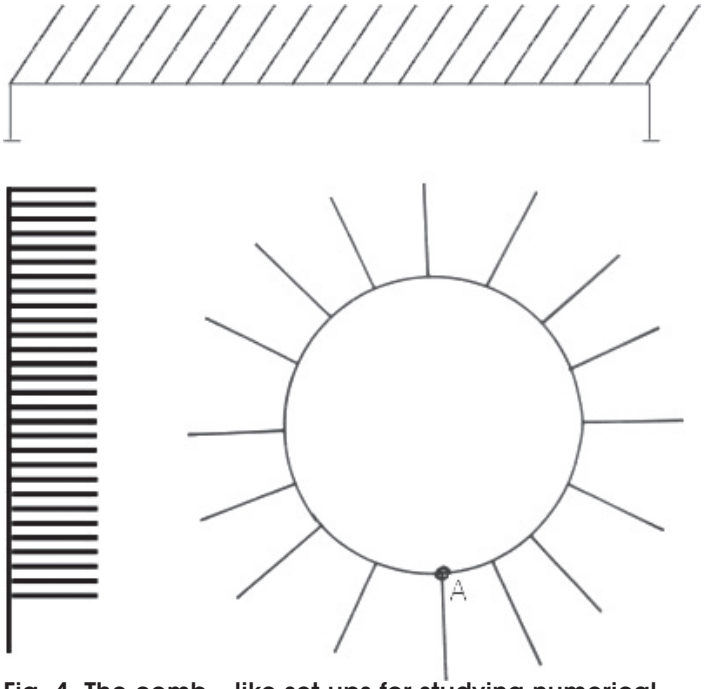


Fig. 4. The comb – like set-ups for studying numerical competence in ants: a horizontal trunk, a vertical trunk and a circle.

The scheme was as follows. Ants were offered a horizontal trunk with 30 branches. The experiments were divided into three stages, and at each of them the regularity of placing the trough with syrup on branches with different numbers was changed. At the first stage, the branch containing the trough with syrup was selected randomly, with equal probabilities for all branches. So the probability of the trough with syrup being placed on a particular branch was 1/30. At the second stage we chose two "special" branches A and B (N 7 and N 14; N 10 and N 20; and N 10 and N 19 in different years) on which the trough with syrup occurred during the experiments much more frequently than on the rest - with a probability of 1/3 for "A" and "B", and 1 /84 for each of the other 28 branches. In this way, two "messages" -"the trough is on branch A" and "the trough is on branch B"- had a much higher probability than the remaining 28 messages. In one series of trials we used only one "special" point A (the branch N 15). On this branch the food appeared with the probability of 1/2, and 1/58 for each of the other 29 branches. At the third stage of the experiment, the number of the branch with the trough was chosen at random again.

The obtained data demonstrated that ants appeared to be forced to develop a new code in order to optimize their messages, and the usage of this new code has to be based on simple arithmetic operations. The patterns of dependence of the information transmission time on the number of the food-containing branch at the first and third stages of experiments were considerably different. In the vicinities of the "special" branches, the time taken for transmission of the information about the number of the branch with the trough was, on the average, shorter. For example, in the first series, at the first stage of the experiments the ants took 70–82 seconds to transmit the information about the fact that the trough with syrup was on branch N 11, and 8–12 seconds to transmit the information about branch N 1. At the third stage it took 5–15 seconds to transmit the information about branch N 11. Analysis of the time duration of information transmission by the ants raises the possibility that at the third stage of the experiment the scouts' messages consisted of two parts: the information about which of the "special" branches was the nearest to the branch with the trough, and the information about how many branches away is the branch with the trough from a certain "special" branch. In other words, the ants, presumably, passed the "name" of the "special" branch nearest to the branch with the trough, and then the number which had to be added or subtracted in order to find the branch with the trough. That ant teams went directly to the "correct" branch enables us to suggest that they performed correctly whatever "mental" operation (subtraction or addition) was to be made (see details in [2]). It is likely that at the third stage of the experiment the ants used simple additions and subtractions, achieving economy in a manner reminiscent of the Roman numeral system when the numbers 10 and 20, 10 and 19 in different series of the experiments, played a role similar to that of the Roman numbers V and X. This also indicates that these insects have a communication system with a great degree of flexibility. Until the frequencies with which the food was placed on different branches started exhibiting regularities, the ants were "encoding" each number (j) of a branch with a message of length proportional to j, which suggests unitary coding. Subsequent changes of code in response to special regularities in the frequencies are in line with a basic information-theoretic principle that in an efficient communication system the frequency of use of a message and the length of that message are related.

In conclusion, we have demonstrated that ants of highly social species can (1) transfer information, (2) compress information, (3) change the way they represent information; (4) add and subtract small numbers. The obtained results are important not only for biology, but also for cognitive science, linguistics, cybernetics and robotics. Generally speaking, we can say that the methods and ideas of Information Theory enabled us to reveal some important laws of Nature.

References

[1] B. Ya. Ryabko and Zh.I. Reznikova, "Using Shannon Entropy and Kolmogorov Complexity to study the communicative system and cognitive capacities in ants", *Complexity*, vol. 2, no.2, pp. 37–42, 1996.

[2] Zh. I. Reznikova, B.Ya. Ryabko, " Numerical competence in animals, with an insight from ants", *Behaviour*, vol. 148, no 4, pp. 405–434, 2011.

[3] Zh. I. Reznikova, *Animal intelligence: From individual to social cognition*, Cambridge University Press, Cambridge, UK, 2007.

[4] E. S. Savage-Rumbaugh, S. G. Shanker, T. J. Taylor, *Apes, Language and the Human Mind*, Oxford University Press, Oxford, UK, 1998.

[5] L. M. Herman, S. L. Abichandani, A. N. Elhajj, E. Y. K. Herman, J. L. Sanchez, A. A. Pack, "Dolphins (*Tursiops truncatus*) comprehend the referential character of the human pointing gesture", *J. Comp. Psychol.*, vol, 113, no. 1–18, 1999.

[6] I. M. Pepperberg, *The Alex Studies,* Harvard University Press, Cambridge, MA, USA, 1999.

[7] B. Ya. Ryabko, Zh. I. Reznikova, "The use of ideas of Informarion Theory for studying language and "intelligence" in ants", *Entropy*, vol. 11, no 4, pp. 836–853, 2009.

Panel on "New Perspectives on Information Theory" IEEE Information Theory Workshop, Paraty, October 20, 2011

Venkat Anantharam (Berkeley)
Giuseppe Caire (USC)
Max Costa (Campinas)
Gerhard Kramer (Technical University Munich)
Raymond Yeung (Chinese University, Hong Kong)
Sergio Verdú, Moderator (Princeton)

A video of the panel discussion (including questions from the audience) can be found in http://media.itsoc.org/itw2011/ What follows is an edited transcript.

Verdú *Seventeen years ago at the ITW that was held in Moscow, I organized a similar panel on the future of Information Theory with the participation of Dick Blahut, Imre Csiszár, Dave Forney, Prakash Narayan and Mark Pinsker. In preparation for this panel I have asked our panelists to read the transcript of that panel (published in the December 1994 issue of this newsletter) and discuss the ways in which that panel's predictions were and were not accurate.*

Costa Well, it´s been said that it is difficult to make predictions, specially about the future. The 1994 panel predictions were good in many aspects, but they could not guess those areas that appeared from nowhere and brought completely new tools and perspectives to the field. There was another situation in which this happened. Estill Green, a VP of Bell Labs, also made some bold and courageous predictions on telecommunications, looking from 1961 into that technology in the year 2012. Bob Lucky commented on those forecasts in 1999, and pointed out the areas in which the forecasts came close to what was happening, and others that were far out. Forecasts usually estimate the increase or reduction of some variables based on the anticipated development of certain known technologies. They are less precise when the actual changes are produced by a complete switch of paradigm, or a totally brand new technology that takes over the field. It was like that that Green's predictions had problems by excluding the changes produced by optical fiber communications and by the advent of the internet. Yet, Green´s predictions were not completely off because he was counting on a very fast growth of video telephony. Bell Labs was investing heavily in video telephony at that time and of course we all know that that project did not go as expected. The idea that an image is worth a thousand words also goes for the rates, and Green was expecting a much greater traffic demand than actually happened from video, but an even greater traffic actually happened for other reasons, like the internet. So it is very difficult to make predictions on a very long stretch of time. Even attempts to update the course of Green's predictions for 2012 only 12 years ago estimated 1 Tbps as the capacity limit of an optical fiber, and we are already seeing lab transmissions of 100 Tbps in a single fiber and 26 Tbps from a single laser source.

Anantharam: There were some really striking omissions. For one thing, there was no mention at all of LDPC codes or turbo codes, which I think has really been one of the defining features of research in our subject for the past decade or so.

Verdú: Actually, the turbo codes had been just presented at ICC, just about a year before, and in fact very few people believed that that paper was correct. And the LDPCs in Gallager's thesis had not been rediscovered yet.

Anantharam: Yes, so that speaks to the point about the danger of predicting the future; another thing to note is that network coding had not yet been invented at the time. From the point of view of someone who is more interested in theoretical

problems, one of the striking things was that some of the problems that were posed as being the major open problems in the field, such as the capacity regions of broadcast channels, relay channels, interference channels–they are all basically still open. Of course, we know a lot more about them and understand them in approximate senses. Also there was a dynamic having to do with what one really means by Information Theory. For instance, Csiszár was one of the panelists and he presented himself as viewing Information Theory more as a mathematical science than a science that is focused entirely on applications. Now, that's not exactly where I stand. I think of Information Theory as a science of understanding how to extract what is relevant about interactions between distributed entities in order for them to coordinate in some way. Communication theory to me is just one aspect of that and in fact much of the discussion in our community is focused on the notion of information in communication theory, but since this panel is meant to look forward, I would venture to say that we are likely to move in directions where the notion of information is coming from fields that are not so close to communication theory, for instance biology, as a prominent example, or neuroscience. So if one is venturing a prediction over a time scale as large as that between 1994 and 2011, I would say that if you are looking at this field 15 to 20 years from now it might not be as communication theory centric as it is now.

Yeung In the 1994 transcript we read that Pinsker and Csiszár were trying to explore the line whether Information Theory is just applied mathematics but also some kind of pure mathematics. My own point of view is that Information Theory is both applied mathematics and pure mathematics, because there is some pure mathematics content in it. In the 1994 transcript, Pinsker brought up the application of Information Theory to ergodic theory and that was a very major impact. Csiszár talked about the application of Information Theory to statistics. This is more like a specific topic as far as I am concerned. But nevertheless, it shows that Information Theory has some pure mathematics elements in it, which is always my belief. In fact I would say that it is this particular reason that has been keeping me very interested in Information Theory. And along this line, during the last 15 years or so, the work by a few others and also by myself on the study of entropy functions has established a link with a few other branches of mathematics. Now we know that there is a common structure between entropy functions, group theory, network coding, and also Kolmogorov complexity. The results along this line also have implications in conditional independence in probability theory. Also, most recently we are interested in how we can use differential entropy functions to study inequalities regarding positive definite matrices. So definitely, I believe that there are some pure mathematics elements in Information Theory. Another thing is that Information Theory is a very special kind of applied mathematics if you want to think of it this way. In many branches of applied mathematics, researchers in the field are users of the results from pure mathematics. You can think of Information Theory being a branch of applied probability, but it is also very different from probability theory, because if you talk to a probability theorist, most likely he or she would know very little about Information Theory. That's why very often in Information Theory we have to develop the tools from scratch. That's something rather unique about Information Theory. Another thing which I find very unique about Information Theory is that for other branches of applied mathematics, we very often use mathematics to model a physical system, for example, control theory, signal analysis, mechanics, and other things. But in Information Theory we are using mathematics to model information, so there is one layer of abstraction, but information by itself is also something rather abstract, instead of being a physical system. So in Information Theory there are two layers of abstraction, which makes the subject very subtle and interesting as far as I am concerned. And therefore after so many years, we still from time to time find that there are some basic things which are not well understood and even not well defined, and that's has been keeping me very interested in the subject.

A few years ago I met a student of Kolmogorov (Albert Shiryaev) who was visiting our university in Hong Kong. When he knew that I was working on Information Theory, he was very interested in talking to me. In fact, Alon Orlitsky was also visiting us at that moment. I told him that I am not a mathematician by training, and so for the mathematics I use all the time I know them quite well, but there are a good number of branches of mathematics that I know nothing about. He said, "That's okay. Shannon was an engineer, too." But he actually was not entirely correct because Shannon got his PhD in mathematics. Then he went on to talk about how Kolmogorov thought about Shannon. I knew that Kolmogorov had worked on Information Theory, but I did not know that Kolmogorov had such a high regard of Shannon. In fact, when Shannon's work was translated to Russian, somehow the section on entropy power inequality was missing. Subsequently, Kolmogorov found out that was the case. He was very upset and said, "How come they could have missed this very important result by Shannon!"

A few years ago (I think it was two years ago) I visited Qualcomm. There I met a Russian guy and we talked about Shannon and Kolmogorov, and then he told me, "You guys in the West compiled a collection of papers by Shannon around 2000. In Russia we had it already in 1963!" Then he rushed to his own office, brought a copy of that collection, and showed it to me. It was all Russian. There I found a word which I thought meant "Shannon." I said, "Is this Shannon's name in Russian?" He said, "Yes." Then he showed me the preface which was written by Kolmogorov. He said, "Basically Kolmogorov said that Shannon had tremendous engineering insight, although he did not really prove anything." That's Kolmogorov's point of view.

Caire What is striking to me while reading the 1994 panel transcripts is the absence of the two keywords that have dominated the last 20 years of communication engineering: internet and wireless. The discussion in 1994 was mainly focused on issues like mathematics vs. engineering, or Shannon theory vs. coding theory. In contrast, the magic word "network" was almost entirely missing. As a matter of fact, network Information Theory has proven to be a formidable area of research both in terms of its richness (the abundance of open problems is almost a life insurance for information theorists . . . . we will never be out of jobs) and in terms of its impact on technology, especially in wireless (3G came in 1996, 4G is happening now, 802.11n with space-time coding, spatial multiplexing happened a few years ago, and the next big thing is to handle interference and multihop relaying). In the wired domain, of course, network coding and its variant application to distributed storage systems (essential for cloud computing) is gaining a lot of traction too.

Kramer I enjoyed the back-and-forth between the engineers (Forney, Blahut), mathematicians (Csiszár, Pinsker), and those perhaps in between (Verdú, Narayan). My apologies if I am categorizing people too much, but our panel does not seem to have the same span of folks as then.

Concerning succeeding and erring on future trends, a few things caught my eye: In contrast to what Giuseppe said, networks were mentioned by several people as an important future topic. Blahut correctly predicted the importance of energy (p. 9) in communications and computing. This is now, 17 years later, a hot topic! Verdú was right in predicting that the probabilistic school is gaining ground on the combinatorial school, in the sense that turbo and LDPC codes dominated coding for many years (with a few exceptions from the then-unknown area of network coding). He also correctly predicted the importance of the interference channel, even if this took a little longer to happen. I think Verdú was too cautious with his statement that "Maybe the day will come when a software package will enable the engineer to closely approach capacity on almost any channel with the technology of the day. Admittedly, I am afraid it is us who will be dead when that day arrives." But about 7 years later, you could download degree distributions for LDPC codes that approach capacity on almost any (practical) channel.

Verdú I am going to stand by the statement I made in 1994. The key is "almost any channel". In fact, even the capacity of many practical single-user channels is still an open problem (e.g. channels with frequency selective and time selective channels, deletion channels). We are nowhere near the point of having an algorithm that given a black box will find not only its capacity but near-optimal codes. Think of the counterpart in data compression, where we do not need to know anything about the source in order to approach its fundamental limits, provided some general technical conditions are satisfied.

In 1994, I did mention the renaissance of physical layer research, but reality ended up exceeding anyone's expectations on the future relevance of Information Theory to the practical world, and in particular the wireless world. The wireless revolution would prove to be a godsend for Information Theory.

Yeung Toby Berger said in the mid-1990's that the rise of wireless is to the advantage of EE (and hence also to IT) because it is something not easily eaten up by CS. He was absolutely right. We have already witnessed the failures of the Google Phone and then Windows Phone.

Verdú I find it interesting that one of the questions I put to the panel in 1994 was "Is Information Theory dead and if not, what evidence do we have to the contrary?" I wouldn't ask that question today: We have come a long way since then and we have now a much higher collective self-esteem and optimism about the future of our field.Why is that?

Kramer One major change since 1992 is the appearance of turbo codes and more generally iterative processing. I like to say that turbo codes made mutual information an engineering quantity. I recall that while I was doing my Master's degree work in 1991, most people believed that that the cut-off rate was the practical limit of reliable communication.

But getting back to the 1994 panel, you had started the discussion with the question "Is Information Theory dead?" and so the focus of the first few pages was of course on this question. I think that the best response was that as long as Information Theory continues to attract the best and brightest young people, Information Theory will continue on as it always has. This response was given by Csiszár (twice! on p. 4, right column, and on p. 10, top left), and by Blahut (p. 11). One can ask what will attract such people, and that will be having difficult and relevant problems (see Forney, p. 9). "Relevant" can mean many things, of course. One thing I especially appreciated from the article was Blahut's statement on p. 5 that it is not necessary to be defensive and negative concerning Information Theory. He wrote that he was living in the "decade of information". And now that decade has stretched on for at least 17 years. I think it will stretch on for many years more. (See Sergio's recent book review on "The Information" in the Sep. issue of the IT Transactions where he laments the focus on 1948.) Blahut was refreshingly positive in his closing statement.

Caire In my opinion, the periodically emerging question about whether IT is dead, or is dying, is essentially due to a complex of inferiority that, at various degrees of intensity, is permeating our community. In fact, I would revert the question and notice that there are other fields, such as, "communications", "signal processing" and "networks", which in the past have been regarded as "eating up" Information Theory, are now getting swallowed by Information Theory.

Today, it is the general understanding of a very broad research and industrial development community that the "Shannon approach" is the winner: extracting the essence of a problem, characterizing its fundamental limits, and designing systems tightly inspired by the optimal or near-optimal strategies stemming from the information theoretic investigation has proven to yield superior results in several areas, and especially in wireless networks. As an example, it is not an accident that systems designed on the basis of engineering common-sense (an infamous example being the 3G CDMA-based system) led to disappointing performance and years of delay in deployment, and that the present generation of networks (LTE, WLANs) is based on OFDM, which is (oversimplifying) what Shannon teaches us to do over a Gaussian frequency selective channel.

If you take a look at 3GPP or IEEE 802.11 standardization forums, it is really striking to notice that what we do percolates almost immediately from theorems to system proposals. In this sense, at least in the realm of wireless communications, Information Theory is the clear winner, "our" approach has become "the approach", and the distance between new discoveries, even the most exotic and probably difficult to implement, such as interference alignment, and the industrial R&D, has become nonexistent. This is also due to Moore's Law: today we can implement on an iPhone 4 algorithms that 20 years ago would have required a powerful mainframe. Such abundant computation power and memory, even in small devices, allows the adoption of Information Theory-driven approaches that in the past were unthinkable.

If a problem exists, right now, is not whether Information Theory is dead or alive, but the lack of recognition that, as a scientific community, we seem to suffer. Our ideas are grabbed very quickly and used by larger and wealthier groups of researchers and

system designers, without always giving the correct credit to the originators. It is sufficient to look at conferences such as SIGCOMM, but sometimes also to ICC and Globecomm, to understand what I am talking about.

Costa Before I address the point, let me make a comment on what Raymond just mentioned. It´s true that some of the proofs of what Shannon could see were not there, like the EPI, for example, the entropy power inequality, that was later proved by Stam and Blachman. But he could have the tremendous insight to see it, and maybe feel that it was essentially the isoperimetric inequality. This amazing insight that Shannon had was also what led him to the random coding argument. I remember a class that Tom Cover gave in a course on Information Theory in which he said: "If I had that idea of the random coding argument, I think I would just go home and sober up." Also to point to the recognition of the work of Shannon, he arrived unexpectedly at the 1985 ISIT in Brighton. Nobody knew that he was going to attend the meeting and there was a big commotion. Bob McEliece is reported to have said that it was just like if in a conference of physicists someone had announced that Newton was present.

Now, we see obituaries of Information Theory come up from time to time. Right now is actually a time that we have a better perspective. Coding theory has also gone through that, and obituaries were announced for coding theory a number of times. One of the early announcements was closely followed by the invention of trellis coded modulation and all the burst of activity that it generated. A few years later, another such obituary was challenged by the creation of turbo codes in 1993, and by the rediscovery of LDPC codes. So it is very risky to make categorical statements regarding the end of an area. Many new things are always coming up to second guess the less optimistic forecaster.

It seems odd to imagine that Information Theory may be perishing when we have just witnessed the dawn of the Information Age. To mention changes that are occurring in a number of schools, the traditional engineering denominations are being replaced by names like information engineering, energy engineering, environmental engineering, and so on. These changes point to the importance of paying attention to the resources, and being resourceful is definitively one of the highlights of Information Theory.

Yeung I just want to add something to what Max just said about Shannon's engineering insight. Well, it has been very amazing to me that while Shannon can be sloppy, there is not a single incidence that he is found wrong. Can someone correct me? It's really amazing! I mean in most mathematical subjects if you do your proofs sloppily, you are bound to get some wrong results, but not in the case of Shannon. I really don't understand.

Verdú I don't think "sloppy" is the right word. The Bell Labs Technical Journal in 1948 is not the IT Transactions in 2011. Shannon wrote a very readable paper aiming to reach a very wide audience. He knew exactly what he was doing. In the paper itself he spelled out a few abstract mathematical arguments, although admittedly he did not include the epsilons and deltas in each and every proof.

Kramer I hope you disagree with me but perhaps the field has become more uniform? For example, at ISITs I think there are fewer pure mathematicians and there are fewer people "directly" connected to industry than before.

Verdú The entropy of topics in ISITs has definitely decreased. Perhaps, today fads play a bigger role in people's choice of fields? I definitely miss the mathematicians, people like Paul Shields, Rudi Ahlswede, Janos Korner, Kati Marton and the whole Russian school. On the positive side, many of the engineers at places like Qualcomm, Bell Labs, IBM, HP Labs are PhDs well-versed in Information Theory.

Caire As a matter of fact, there have been years where an enormous number of sessions was focused on just one topic, and referenced only a handful of papers (e.g., at the peak of Trellis Coded Modulation, the topic distribution was essentially a delta function at Ungerboeck). Right now, we have topics such as biology, compressed sensing and matrix reconstruction, wireless networks, network coding, machine learning, and many more.

Verdú Going back to answering my question on the renaissance of Information Theory, I think that much of this renaissance comes from: 1) the great influx of young talent in the last two decades; 2) the successes of information-theory-driven technologies such as: sparse-graph codes, universal data compression, voiceband modems, discrete multitone modulation (DSL), multiuser detection and MIMO (Shannon tells us don't treat digitally modulated signals as thermal noise: exploit their structure), space-time codes, opportunistic signaling, network coding, etc. I believe that there is increasing realization in industry that Information Theory provides the only reliable guidance for sound efficient design. It used to be that technology was way behind theory, and the big challenge was "how to do it?". Often now the bottleneck is not implementation but lack of complete theoretical understanding. The challenge is "what to do?"

In this respect, it is useful to contrast IT with the Complexity Theory community within Theoretical Computer Science: a relatively small community of abstract thinkers occupied with fundamental limits of efficient computation. The holy grail NP $\neq$ P seems more elusive than ever. Lately they are devoting a lot of efforts to studying the role of randomness in computation, the foundations of cryptography, interactive proofs, quantum computational complexity. A leading member of that community, Andy Yao, describes the situation as "a bunch of monkeys climbing trees in order to reach the moon." The big difference is that the real world of technology keeps us honest and relevant. That gives us a lot of credibility for the outside world. For us a 10% improvement grabs our attention, for the complexity theory people there is little difference between linear and n^{1000}.

Anantharam I might like to add a couple of notes of caution while we are busy congratulating ourselves on how successful we have been. I think we have also been great beneficiaries of a lot of serendipity. My own view of research is that there is a certain kind of randomness involved in the generation of new ideas and many of the ideas that have driven the field are not really of our own doing. There are great individual results that have come out perhaps with atypical frequency over the last couple of decades which one could not really have predicted. Quite apart from that, there is also the general evolution of technology. Moore's law, as Giuseppe mentioned, the driver that hardware provides to create problems where we are relevant is something that we shouldn't forget. In fact, one can mention the possible applications of some of our techniques at levels that were not even conceived of at the time of the previous panel, for instance

coding at the level of communication over VLSI buses. So my main point in the context of this discussion is that we should be a little bit humble about why we are so successful—it's not entirely of our own making.

Verdú *Just like in 1994 I think it is futile to try to predict what the disruptive problems that will revolutionize the field in the future. Nevertheless, it is useful to discuss those current topics with the highest chance to have an impact on the future development of Information Theory.*

Kramer I suppose we all have our favorite current topics that we feel are important now and that we can predict will be important in the future. One of my favorites is Information Theory applied to optical channels, including optical fiber (MIMO is hot), free space, non-coherent vs. coherent, quantum, and so forth. A second favorite topic of mine is whether and how one can transfer the substantial progress on understanding relay and interference channel capacity into wireless systems. A third favorite is the same question concerning network coding and its application to distributed storage.

Caire At the risk of being disproved and laughed about by those who will read the transcript of this panel in a few years from now, I am going to try a forecast. What I'd like to see in the next 5–6 years from now is the development of a "communication theory for networks". We gained an enormous insight about network Information Theory, in understanding interference and relaying. Nevertheless, we are very far from a "plug-and-play" set of techniques around which novel physical layer architectures can be actually designed. To make an analogy, in point-to-point communications we had Ungerboeck TCM, and now Turbo and LDPC codes followed by some form of bit-interleaving and mapping onto modulation alphabets. These techniques have been widely studied at the point that they have become standard tools around which systems based on point-to-point links can be safely designed. Still, in order to handle interference we rely on orthogonal (or quasi-orthogonal) access, and treating interference as noise, or as "collisions". We are still very far from the point where a new set of codes in the signal space (lattice codes? polar codes applied to multiuser problems?) can be used as basic building blocks for a robust system design. Of course, the risk of not filling this gap between Information Theory and communication theory (and therefore, practice of system design) is that these areas will remain confined in the purely theoretical domain and they may eventually fade away.

Anantharam As I said in the beginning, much of the success we have had in this field is centered around problems that are in the communication theory arena, but I think there are vast realms out there that are waiting to be conquered by what you might call "information-theoretic thinking". Shannon basically brought information-theoretic thinking to bear on communication theory. But there are aspects of nature, for instance, which have to have been designed with the concept of the optimization of some kind of information content in view. When you have a lot of interacting entities, either entities in nature or entities that you want to design, for instance when you want to design a biological system, which is eventually going to happen, there has to be a thinking, both in the engineered design and in nature as it came up with the designs that we are aware of today, which involves a notion of some information aspect that was optimized in enabling the coordination between the interacting entities. I am not sure on what time scale this will happen, but for instance biology is advancing at an enormous rate, so it could very well be twenty years, maybe longer, I think what we will see is a success of information-theoretic thinking in a lot of fields that have nothing to do with communication theory. Of course we are going to see all these great things happening in communication theory, which is very close to our hearts, but if you were to come to a meeting with the title "Information Theory workshop" twenty years from now, maybe we will have, say 20% of the papers where people are discussing for instance why a certain organ in the body works the way it does because the cells have enabled the coordination between themselves by optimizing some kind of information measure. We are really waiting for the Shannon for all of these different fields. That Shannon hasn't arrived and it is not clear if that Information Theory will look the way Shannon's Information Theory does, but I think we can rest assured that the Shannon will eventually come.

Verdú We are waiting for the second coming of the Messiah.

Costa I mentioned something about the increase in capacity that we have already seen, with fibers transmitting at 100 Tbps. But just to mention something that is anticipated on the demand side, the traffic in the internet is supposed to multiply by four in the next four years. We are supposed to have 15 billion network connected devices by 2015, and the expected overall traffic is supposed to be one zetabyte (10^{21}) per year in 2015. So when we get together for ISIT in Hong Kong in 2015, we will be able to check on these predictions.

In the long run, not focusing strictly in communications, but in the broader aspects of Information Theory, I think there will be a number of changes that will come up, particularly because of some of the connections that were already mentioned, with economics and biology. In fact we have already seen some BCH code structures in some proteins, and I think that the secrets of biological codes will be revealed little by little, protein by protein. Of course that will have a tremendous impact on what we will be able to do. I think polar codes will also have a great impact, and they will be extended to multiple user channels. We have already seen some of this happening with multiple access channels. Modern coding theory is now basically prevalent and to some extent has replaced the drive in algebraic coding theory, but I think that algebraic coding theory will make a strong come back in network coding applications and in the design of cyclic and quasi-cyclic LDPC codes with more predictable performances and substantially decreased error floors. Also quantum codes will come to be something that people will relate to in a more pragmatic way. And their anticipated impact, I believe, is enormous.

Another thing that I notice is that I remember Tom Cover, many years ago, talking about multiple user Information Theory as the foundation to a more general network understanding. Of course the setting at that time was completely different. (Incidentally, again Shannon was the first one to write something about multiple user channels with his two-way channel paper in 1961.) Network coding has brought techniques to surpass the classical max-flow-min-cut limits, and even the famous butterfly network has seen improvements in some cases, when the intermediate router does not need to access all the inputs, but needs just a function of these inputs. Even though we have had these tremendous advances in network coding theory we have not yet seen a marriage of that theory with the basic

tiles or bricks that form multiple user structures, like broadcast channels, multiple access channels, interference channels and relay channels. I think there will be more development in these areas, and it may be a stretch, it may take a long time, but eventually we will see some conciliation and integration between the approaches of network coding and multiple user Information Theory.

Also source coding is still far from achieving the limits. I must say that I didn't expect to be alive on the day that channel coding limits would be approached as they are, within a small fraction of a dB. I really thought that this would be happening after my time. Now we can ponder that source coding is still not at that point, and hopefully we will still be around when those limits are approached within a fraction of a dB. More effective ways to combine source and channel coding will also became prominent. Improved and new inequalities will continue to extend the power and the scope of Information Theory tools.

I believe there are many fronts in which Information Theory will continue to bring significant contributions, both in practical technologies and in pure scientific and mathematical issues. So rather than thinking that Information Theory will eventually come to some type of blockade, I am more inclined to think that Information Theory will never die.

I would like to quote Karl Popper on something similar to the idea of monkeys trying to reach the moon. My son Bruno is a philosopher and we have some interesting discussions about this sort of thing. This is something that he told me. Karl Popper used to say that we may be very different in the ways we do things and in our abilities and knowledge about things, but in our infinite ignorance we are all alike.

Yeung A few years ago I had a chat with Prakash Narayan who was a panelist 17 years ago. I was pointing out the fact in the control theory community people had been using optimization tools for decades. Prakash made a very interesting point. He said that once the structure of a problem is exposed, what remain are algorithms and optimization. So, I think at least in the context of communications, in the Information Theory community we are going to see more and more of that.

Since we are still on the broad topic of "New Perspectives for Information Theory," I want to pick up a point Sergio mentioned a little while ago, regarding the lack of entropy of topics at ISIT. Personally, this actually bothers me quite a bit. I remember visiting Jim Massey in 2000 at Copenhagen. I was mentioning to Jim that these days research has become so competitive in many areas that if you don't publish something immediately, then very likely 3 months later somebody else will publish the same result. And Jim said, "In that case you shouldn't publish the result." I am not sure whether this is the best way to survive in today's research environment. Ideally we should all be working on problems that we think are important instead of following what the trend is doing all the time. But in the United States in particular, research is pretty much driven by funding. Whatever they call for you have to work on it, although you can do things in disguise. You have to follow the game.

Verdú Looking at my crystal ball, I see going forward: breakthroughs in multiuser Information Theory; intersections of Information Theory with machine learning, with signal processing, with compressed sensing, with theoretical computer science. Maybe one day we will think of the beginning of the XXI century as the era of bad quality: dropped cellphone calls, bad skype connections, lousy youtube video quality. This should put pressure in narrowing the gap between lossy compression theory and practice, and to that end one of the requisites is to learn how we can fool the eye and the ear more effectively than today. New approaches drawn from other fields such as random matrices and statistical physics methods are gaining prominence. And finally, non-asymptotic Information Theory: many practical applications are characterized by short messages or strict delay constraints. In the non-asymptotic regime we do not have the luxury of the closed-form formula, but we can still get very tight bounds as a function of delay.

Anantharam Other modern mathematical tools are also being brought to bear on Information Theory problems, e.g. from additive number theory in problems of interference alignment, and new kinds of concentration inequalities from our improved understanding of concentration of measure. Extracting structure from randomness is central to many branches of mathematics.

Verdú *It is time for me to thank all the attendees, my fellow panelists, and Valdemar da Rocha and Sueli Costa for organizing this workshop and providing the impetus for the organization of this panel.*

Addendum *Ioannis Kontoyiannis of the Athens University of Economics and Business was scheduled to participate in the panel but had to cancel his appearance. Here are some of his thoughts regarding the panel discussion.*

Kontoyiannis Reading the transcript of the panel discussion that was held at ITW in 1994, one notices that our community has made absolutely no progress in answering "foundational" questions like, "Is Information Theory part of applied mathematics or is it an engineering discipline?" I consider this a great success. About 15 years ago, the speaker in a philosophy of science seminar I was attending remarked that, when a field enters existential, esoteric discussions of this kind, it is usually a sign of intellectual decline. My (admittedly self-serving) view is that, as a field, we have been so successful that we can afford to avoid entertaining these questions seriously. When things are looking up, one rarely worries about the meaning of life. This success, as far as I can judge, has been facilitated to a significant extent by the combination of two distinct qualities. The field of Information Theory is defined by basic problems we wish to solve; and the community is fearless in bringing in the right tools to attack these problems. We are not a "methods looking for problems" discipline, and we have been open to the use of whatever new mathematical tool works – from the traditional analytical and probabilistic machinery of applied mathematics to the use of elliptic curves, random matrices, additive combinatorics, and so on.

On the other hand, the field is mature enough that it is now seen by researchers in numerous other areas as a collection of useful tools for their problems. The near simultaneous appearance of special issues on "Information Theory in neuroscience" in our Transactions and in the Journal of Computational Neuroscience is strong evidence of this trend. The editorial in the JCN special issue concludes: "Information Theory is thriving in the neuroscience community, and the long efforts are bearing fruit, as diverse research questions are being approached with more elaborate

and refined tools. […] Information Theory is firmly integrated in the fabric of neuroscience research, and a progressively wider range of biological research in general, and will continue to play an important role in these disciplines."

Shannon's bandwagon warning notwithstanding, it is probably a safe prediction that this trend – information theoretic-ideas and tools being systematically applied in biology and perhaps in the other sciences – will continue and it will grow. In the reverse direction, another recent –though somewhat less noticeable– trend has been the growing use of information-theoretic concepts in core mathematics research. Although this was advocated by Kolmogorov almost 30 years ago ("Information Theory must precede probability theory and not be based on it. [...] The concepts of Information Theory […] can acquire a certain value in the investigation of the algorithmic side of mathematics as a whole"), progress has perhaps been slower and less flashy than the corresponding successes in, e.g., biology. But there are numerous examples – including Perelman's proof of the Poincaré conjecture and the celebrated Green-Tao theorem on the existence of arithmetic progressions in the primes – where Shannon entropy and the associated "technology" have served as important intellectual guidelines for major mathematical breakthroughs. This is another direction that I believe will continue strong and will gain momentum.

Finally, one of the essential components of our trade has to do with building foundations. Given a new communications scenario – be it a new technology with different physical characteristics, a new biological setting describing the communication between two distinct parts of an organism, or a new type of network model like those we have been studying in recent years arising in social media interactions – we abstract its fundamental characteristics and provide a rigorous "language" for its study. Keeping an open mind – and open doors – towards such new problems virtually guarantees a healthy outlook and a wealth of opportunities. A recent success story in this direction is the area of "compressed sensing." This could well have become a sub-field of statistics or harmonic analysis. The fact that it was embraced by the Information Theory community is a testament to both our open-mindedness and our strength.

I cannot resist one last comment. We really need to figure out how to do lossy compression effectively in practice!

Report on the Princeton CCI Workshop on Counting, Inference, and Optimization on Graphs

Jin-Yi Cai, Michael Chertkov, G. David Forney, Jr., Pascal O. Vontobel, and Martin J. Wainwright (co-organizers)

Over 100 participants attended an interdisciplinary workshop on "Counting, Inference, and Optimization on Graphs" at Princeton University, NJ, November 2–5, 2011. The workshop was organized by the authors under the auspices of the Center for Computational Intractability (CCI) at Princeton.

The workshop was originally inspired by the recognition of connections between certain duality results in the theory of codes on graphs and recent work on "holographic" algorithms in theoretical computer science. Ultimately, topics included holographic algorithms, complexity dichotomy theorems, capacity of constrained codes, graphical models and iterative decoding algorithms, and exact and approximate calculation of partition functions of graphical models. The participants had a wide range of backgrounds, including theoretical computer science, information and coding theory, statistical physics, and statistical inference.

The program is listed below. Copies of slides, references to related papers, and videos of some of the talks are available on the conference website at <http://intractability.princeton.edu/blog/2011/05/workshop-counting-inference-and-optimization-on-graphs>.

The participants were enthusiastic about the quality of the talks, the stimulation of various cross-disciplinary dialogues, and the excellent arrangements provided by the Center for Computational Intractability.

Program:

Leslie Valiant, "Holographic algorithms"

Jin-Yi Cai, "Complexity dichotomy theorems for counting problems"

Mark Jerrum, "Approximating the partition function of the ferromagnetic Ising model"

Leslie Ann Goldberg, "Approximating the Tutte polynomial (and the Potts partition function)"

Martin Loebl, "Complexity of graph polynomials"

Predrag Cvitanović, "Dynamical zeta functions: What, why, and what are they good for?"

Michael Chertkov, "Gauge transformations and loop calculus: General theory and applications to permanents"

Moshe Schwartz, "Networks of relations in the service of constrained coding"

Vladimir Chernyak, "Planar and surface graphical models which are easy"

Farzad Parvaresh, "Asymptotic enumeration of binary matrices with bounded row and column weights"

David Forney, "Codes on graphs, normal realizations, and partition functions"

Yongyi Mao, "Normal factor graphs, linear algebra and probabilistic modeling"

Navin Kashyap, "The tree width of a linear code"

Pascal Vontobel, "Should we believe in numbers computed by loopy belief propagation?"

Martin Dyer, "On the complexity of #CSP"

Xi Chen, "Complexity of counting CSP with complex weights"

Alistair Sinclair, "Permanents, determinants, and commutativity"

Leonid Gurvits, "A new entries-dependent lower bound on the permanent of doubly stochastic matrices"

Martin Wainwright, "Learning in graphical models: Missing data and rigorous guarantees with non-convexity"

Yair Weiss, "Convexity: What is it good for?"

Marc Mézard, "Statistical physics-based reconstruction in compressed sensing"

Andrea Montanari, "Sharp thresholds in statistical learning"

Anima Anandkumar, "High-dimensional graphical model selection: Tractable graph families and regimes"

Umesh Vazirani, "Quantum description complexity"

David Poulin, "Belief propagation in the quantum world"

Jonathan Yedidia, "The alternating direction method of multipliers as a message-passing algorithm"

Alexander Ihler, "Variational algorithms for marginal MAP"

Alexander Barvinok, "Counting integer points in polyhedra"

David Gamarnik, "Interpolation method and scaling limits in sparse random graphs"

Mehdi Molkaraie, "Monte Carlo algorithms for the partition function of two-dimensional channels"

Pinyan Lu, "Approximate counting via correlation decay in spin systems"

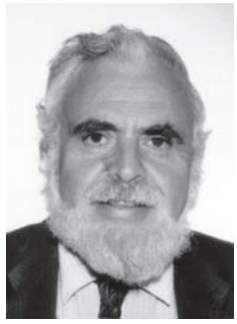
Jinwoo Shin, "Improved mixing condition on the grid for counting and sampling independent sets"

President's Column *continued from page 1*

Networking, most computing devices are used more for communicating than for computing *per se*. Further afield from some of our traditional collaborations, areas such as computational biology present great promise for activities in our Society. The information age surely must benefit broadly from information theory. It is my hope that, as we ponder the direction of our Society, we consider highlighting past, existing and emerging contributions of our Society to various areas, such as computer science. Board of Governors member Michelle Effros is currently leading and *ad-hoc* committee, which was formed by Giuseppe to look into potential opportunities for our Society to advocate activities. Of course, intellectual benefits are mutual. Our Society I believe will remain open to contributions from, as well as to, different areas, and will continue to value applications not only as mere instantiations of theory but as a valuable source of ideas. It is my sincere hope to be able to help, as your president, in furthering the conversation within our Society, as well as with the many technical communities with which we intersect, about our future directions.

GOLOMB'S PUZZLE COLUMN™

Powers with Shared Digits



Solomon W. Golomb

A positive integer which is a square, cube, or higher power of some integer will be called simply a *power*. The infinite sequence of powers begins $\{1,4,8,9,16,25,27,32,36,49,64,81,100,121,125,128,\ldots\}$.

1) How many powers are there from 1 to 1 million?

2) How many powers are there from 1 to x, for real $x > 4$?

Sometimes, two or more powers will consist of the same k digits, though in different permuted orders. As seen above, there are no such cases with $k = 2$ digits.

3) There are several cases with powers having the same $k = 3$ digits. Try to find these:

   a) Three different powers, all squares, with the same three distinct digits.

   b) Two different cubes with the same three distinct digits.

   c) Two squares with the same three digits, where two of these three digits are the same.

   d) A square and a cube share the same three distinct digits, but in different orders.

   e) Two different fourth powers share the same three digits. (*Note*. We do not allow powers to have "0" as their most significant digit. Thus, $3^2 = 009, 30^2 = 900,$, is not an acceptable solution to 3.c.)

4) Here are some sets of powers that share the same $k = 4$ digits. Can you find these?

   a) Three different four-digit numbers, a square, a cube, and a fourth power, have the same four distinct non-zero digits.

   b) Another pair of squares share a different set of four distinct non-zero digits.

   c) Two more pairs of squares each share their own set of four distinct digits. (These sets include the digit "0", but always in an intermediate position.) Both members of one of these pairs are in fact even powers higher than the second power.

5) There is a unique set of five distinct non-zero digits whose $5! = 120$ permuted orders include five different squares, two of which are also higher powers than squares. (This fact may help you find these, since even powers above squares are relatively rare.)

6) Two four-digit powers, **a** and **b**, which share the same distinct digits (each is an even power higher than a square), have squares $\boldsymbol{a}^2$ and $\boldsymbol{b}^2$, which share the same set of seven distinct digits. Can you find this remarkable example?

GOLOMB'S PUZZLE COLUMN™

The Sequence n^3-n Solutions

Solomon W. Golomb

1) The only cases where $n^3 - n = k!$ are: $n = 2, k = 3; n = 3, k = 4; n = 5, k = 5; n = 9, k = 6$.

2) Since $n^3 - n = (n-1)n(n+1)$ is a product of three consecutive integers, one of these factors is a multiple of 3, and at least one of them is a multiple of 2. For $n > 3$, if $n^3 - n$ has only two different prime factors, all three of $n-1, n$, and $n+1$ must be divisible only by 2 and 3. This occurs with $n = 3$, where $3^3 - 3 = 2 \cdot 3 \cdot 4 = 2^3 \cdot 3$. For $n > 3$, if n is even, both $n-1$ and $n+1$ are odd and ≥ 3, and only one of them can be a power of 3, so the other has a prime factor which is neither 2 nor 3. If n is odd, both $n-1$ and $n+1$ are even, and only one of them can be divisible by $4 = 2^2$. Since $n > 3$, the other one has a prime factor q other than 2. If $q = 3, n$ is divisible by neither 2 nor 3, and contains a third prime factor. If $q \neq 3$, then 2, 3, q are three different prime factors of $n^3 - n$.

3) The values of n for which $n^3 - n$ has exactly three distinct prime factors are $n = 4, 5, 7, 8, 9$, and 17.

4) The three situations that make it possible for $n^3 - n$ to have (only) four different prime factors are that one of $n-1, n$, or $n+1$ is a) a power of 2, b) a power of 3, or c) a power of 2 times a power of 3. (It is necessary, but not sufficient, that one of these three conditions be satisfied.)

5) These are the 63 values of $n, 3 < n < 10^6$, for which $n^3 - n$ has exactly four different prime factors:
6, 10, 11, 12, 13, 15, 16, 18, 19, 23, 24, 25, 26, 27, 28, 31, 32, 33, 37, 47, 48, 49, 53, 63, 72, 73, 80, 81, 82, 87, 107, 108, 127, 128, 163, 192, 193, 242, 243, 257, 283, 432, 487, 513, 577, 863, 1152, 1153, 2187, 2592, 2593, 2917, 4373, 8192, 8747, 131072, 131073, 139968, 472392, 524288, 786432, 995327, 995328.

6) As in 2. above, for all $n \geq 2, n^3 - n$ must be divisible by both 2 and 3, where 3 may divide any one (and only one) of $n-1, n$ or $n+1$; but 2 (and not 2^2) must divide n, so we must have $n = 4a + 2$ for some $a \geq 0$.

7) The fraction of integers, asymptotically, for which $n^3 - n$ is square-free equals

$$\prod_{\text{all } p} \left(1 - \frac{3}{p^2}\right) = 0.12548698\ldots,$$

where the product is taken over all prime numbers p.

8) When $n^3 - n$ is square-free, we have $n = 4a + 2$. Then, 3 divides one of $n, n-1$, or $n+1$. If 3 divides $n = 4a + 2$, and if $n > 6$, then n has a prime factor different from both 2 and 3, while $n-1$ and $n+1$ each contribute at least one additional prime factor, for a total of at least 5 distinct prime factors of $n^3 - n$. If 3 divides $n-1$ and $n > 4$, then each of $n-1, n$, and $n+1$ contribute at least one additional prime factor, other than 2 or 3, for a total of at least five distinct prime factors of $n^3 - n$. Similarly, if 3 divides $n+1$ and $n > 2$, then each of $n-1, n$, and $n+1$ contribute at least one additional prime factor, other than 2 or 3, for a total of at least five distinct prime factors of $n^3 - n$.

9) The only values of n where $n^3 - n$ is square-free with $k < 5$ distinct prime factors are: $n = 2$, where $2^3 - 2 = 6 = 2 \cdot 3$ has $k = 2$ distinct prime factors; and $n = 6$, where $6^3 - 6 = 210 = 2 \cdot 3 \cdot 5 \cdot 7$ has $k = 4$ distinct prime factors.

10) Here are the first three occurrences of n where $n^3 - n$ is square-free with k distinct prime factors, for each $k, 5 \leq k \leq 10$.

$k = 5, n = 14, 22, 30.$	$k = 8, n = 286, 610, 806.$
$k = 6, n = 34, 66, 70.$	$k = 9, n = 714, 1310, 1770.$
$k = 7, n = 186, 210, 230.$	$k = 10, n = 7314, 7566, 7734$

Here are the factorizations.

$k = 5.$
$14^3 - 14 = 2 \cdot 3 \cdot 5 \cdot 7 \cdot 13$
$22^3 - 22 = 2 \cdot 3 \cdot 7 \cdot 11 \cdot 23$
$30^3 - 30 = 2 \cdot 3 \cdot 5 \cdot 29 \cdot 31$

$k = 6.$
$34^3 - 34 = 2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdot 17$
$66^3 - 66 = 2 \cdot 3 \cdot 5 \cdot 11 \cdot 13 \cdot 67$
$70^3 - 70 = 2 \cdot 3 \cdot 5 \cdot 7 \cdot 23 \cdot 71$

$k = 7.$
$186^3 - 186 = 2 \cdot 3 \cdot 5 \cdot 11 \cdot 17 \cdot 31 \cdot 37$
$210^3 - 210 = 2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdot 19 \cdot 211$
$230^3 - 230 = 2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdot 23 \cdot 229.$

$k = 8.$
$286^3 - 286 = 2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdot 13 \cdot 19 \cdot 41$
$610^3 - 610 = 2 \cdot 3 \cdot 5 \cdot 7 \cdot 13 \cdot 29 \cdot 47 \cdot 61$
$806^3 - 806 = 2 \cdot 3 \cdot 5 \cdot 7 \cdot 13 \cdot 23 \cdot 31 \cdot 269$

$k = 9.$
$714^3 - 714 = 2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdot 13 \cdot 17 \cdot 23 \cdot 31$
$1310^3 - 1310 = 2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdot 17 \cdot 19 \cdot 23 \cdot 131$
$1770^3 - 1770 = 2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdot 23 \cdot 29 \cdot 59 \cdot 61$

$k = 10.$
$7314^3 - 7314 = 2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdot 19 \cdot 23 \cdot 53 \cdot 71 \cdot 103$
$7566^3 - 7566 = 2 \cdot 3 \cdot 5 \cdot 7 \cdot 13 \cdot 17 \cdot 23 \cdot 47 \cdot 89 \cdot 97$
$7734^3 - 7734 = 2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdot 13 \cdot 17 \cdot 19 \cdot 37 \cdot 1289$

Call for Nominations

IEEE Information Theory Society 2012 Claude E. Shannon Award

The IEEE Information Theory Society Claude E. Shannon Award is given annually for consistent and profound contributions to the field of information theory. Award winners are expected to deliver the Shannon Lecture at the annual IEEE International Symposium on Information Theory held in the year of the award.

NOMINATION PROCEDURE: Nominations and letters of endorsement must be submitted by March 1, 2012 to the current President of the IEEE Information Theory Society, who in 2012 will be Muriel Medard <medard@MIT.edu>. The nomination form is available at http://www.itsoc.org/honors/claude-e.-shannon-award.

(See page 35 for more information).

IEEE Information Theory Society 2012 Aaron D. Wyner Distinguished Service Award

The IT Society Aaron D. Wyner Award honors individuals who have shown outstanding leadership in, and provided long standing exceptional service to, the Information Theory community. Each Wyner Award winner receives an ISIT or ITW participation fee waiver, a specially engraved plague, and a certificate. This award was formerly known as the IT Society Distinguished Service Award.

NOMINATION PROCEDURE: Nominations and letters of endorsement must be submitted by March 1, 2012 to the current President of the IEEE Information Theory Society, who in 2012 will be Muriel Medard <medard@MIT.edu>. The nomination form is available at http://www.itsoc.org/honors/wyner.

(See page 35 for more information).

IEEE Information Theory Society 2012 Paper Award

The Information Theory Society Paper Award is given annually for an outstanding publication in the fields of interest to the Society appearing anywhere during the preceding two calendar years (2010–2011). The purpose of this 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 Award consists of a certificate and an honorarium of US$1,000 for a paper with a single author, or US$2,000 equally split among multiple authors. The award will be given for a paper published in the two preceding years.

NOMINATION PROCEDURE: Nominations and optional letters of endorsement must be submitted by March 15, 2012 to the Awards Committee chair, who in 2012 will be Gerhard Kramer <gerhard.kramer@tum.de>. Please email the name of the paper you wish to nominate, along with a supporting statement explaining its contributions.

IEEE Joint Comsoc/IT 2012 Paper Award

The Joint Communications Society/Information Theory Society Paper Award recognizes outstanding papers that lie at the intersection of communications and information theory. Any paper appearing in a ComSoc or IT Society publication during the preceding three calendar years (2009–2011) is eligible for the 2012 award. A Committee with members from both societies will make the selection. The award consists of a plaque and cash prize presented at the Comsoc or IT symposium of the authors' chosing.

NOMINATION PROCEDURE: By February 15, 2012, please email the name of the paper you wish to nominate, along with a supporting statement explaining its contributions to both communications and information theory, to the Awards Committee chair, who in 2012 will be Gerhard Kramer <gerhard.kramer@tum.de>.

IEEE Fellow Program

For (s)he's a jolly good (IEEE) Fellow! Do you have a friend or colleague who is a senior member of IEEE and is deserving of election to IEEE Fellow status? If so, consider submitting a nomination on his or her behalf to the IEEE Fellow Committee. The deadline for nominations is March 1st. IEEE Fellow status is granted to a person with an extraordinary record of accomplishments. The honor is conferred by the IEEE Board of Directors, and the total number of elected Fellows in any one year is limited to 0.1% of the IEEE voting membership. For further details on the nomination process please consult: http://www.ieee.org/web/membership/fellows/index.html.

IEEE Awards

The IEEE Awards program has paid tribute to technical professionals whose exceptional achievements and outstanding contributions have made a lasting impact on technology, society and the engineering profession. For information on the Awards program, and for nomination procedures, please refer to http://www.ieee.org/portal/pages/about/awards/index.html.

2012 IEEE European School of Information Theory April 16-20, 2012, Antalya, Turkey

http://www.itsoc.org/european-school

The 2012 IEEE European School of Information Theory will take place in Antalya, Turkey between the 16th and the 20th of April, 2012. The event, organized jointly by CTTC (Spain), TUM (Germany) and Bahcesehir University (Turkey), will offer graduate students and young researchers the opportunity to learn from experts in information theory through half-day tutorials, as well as the chance to present and discuss their own ongoing work.

This is the 12th information theory school in Europe, and we again have a distinguished list of speakers. This year's instructors and tentative lecture titles are:

* Frans Willems (Eindhoven University of Technology, Netherlands) Introduction to Universal Source Coding and Biometrics

* Sennur Ulukus (University of Maryland, USA) Information Theoretic Security

* Meir Feder (Tel Aviv University, Israel) Efficient Lattice Codes

* Alex Dimakis (University of Southern California, USA) Network Coding for Distributed Storage

* Michael Gastpar (UC Berkeley, USA and EPFL, Switzerland) Algebraic Structure in Network Information Theory

The IEEE Information Theory Society is the main sponsor of the 2012 European School of Information Theory.

General Chairs: Deniz Gunduz (CTTC), Gerhard Kramer (TUM)

Local Organization Chair: Alkan Soysal (Bahcesehir University)

For additional information, see: http://www.itsoc.org/european-school

FIFTIETH ANNUAL ALLERTON CONFERENCE

ON COMMUNICATION, CONTROL, AND COMPUTING

October 1 – 5, 2012
Call for Papers

The Fiftieth Annual Allerton Conference on Communication, Control, and Computing will be held from Monday, October 1 through Friday, October 5, 2012, at Allerton House, the conference center of the University of Illinois. Allerton House is located twenty-six miles southwest of the Urbana-Champaign campus of the University in a wooded area on the Sangamon River. It is part of the fifteen-hundred acre Robert Allerton Park, a complex of natural and man-made beauty designated as a National natural landmark. Allerton Park has twenty miles of well-maintained trails and a living gallery of formal gardens, studded with sculptures collected from around the world.

Papers presenting original research are solicited in the areas of communication systems, communication and computer networks, detection and estimation theory, information theory, error control coding, source coding and data compression, network algorithms, control systems, robust and nonlinear control, adaptive control, optimization, dynamic games, multi-agent systems, large-scale systems, robotics and automation, manufacturing systems, discrete event systems, multivariable control, computer vision-based control, learning theory, cyber-physical systems, security and resilience in networks, VLSI architectures for communications and signal processing, and intelligent transportation systems.

Allerton Conference will be celebrating its Golden Anniversary this year. Because of this special occasion, the conference will be longer than its usual 2 ½ days format, to accommodate some special sessions and events in connection with the 50th celebration.

Information for authors: Regular papers suitable for presentation in twenty minutes are solicited. Regular papers will be published in full (subject to a maximum length of eight 8.5" x 11" pages, in two column format) in the Conference Proceedings. Only papers that are actually presented at the conference can be included in the proceedings, which will be available after the conference on IEEE Xplore.

For reviewing purposes of papers, a title and a five to ten page extended abstract, including references and sufficient detail to permit careful reviewing, are required.

Manuscripts must be submitted by **Tuesday, July 10, 2012**, following the instructions at the Conference website: http://www.csl.uiuc.edu/allerton/**.**

Authors will be notified of acceptance via e-mail by August 3, 2012, at which time they will also be sent detailed instructions for the preparation of their papers for the Proceedings.

Final versions of papers to be presented at the conference will need to be submitted electronically by October 5, 2012.

Conference Co-Chairs: Bruce Hajek and Tamer Başar
Email: *allerton@csl.uiuc.edu* URL: http://www.csl.uiuc.edu/allerton

COORDINATED SCIENCE LABORATORY AND THE DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

University of Illinois at Urbana-Champaign

2012 IEEE COMMUNICATIONS THEORY WORKSHOP

Ka'anapali, Maui, Hawaii, USA
May 14-16, 2012

Call for Posters

The technical program committee is soliciting contributions for early-evening poster sessions devoted to recent results. We especially encourage students and junior researchers to participate. There will be no published proceedings. Any contribution in the general area of communication theory is welcome. Extended abstracts (2 pages maximum, double column, IEEE style file) for the poster session should be submitted by March 1, 2012. Electronic submissions must be in Adobe PDF format, and of sufficient detail to permit careful reviewing. Acceptance notifications will be sent on March 22, 2012. The advance-registration deadline is April 5, 2012.

Organizing Committee

Co-chairs:
Randall Berry, Michael Honig

TPC chair:
Venu Veeravalli

Finance chair:
Dongning Guo

Posters chair:
Vijay Subramanian

Industrial Liasons:
Ian Collings, David Love

Publicity chairs:
Natasha Devroye, Jianwei Huang

Advisory Board:
Jeffrey Andrews, Andrea Goldsmith, Robert Heath, Michael Rice, Reinaldo Valenzuela

Sponsored by:

IEEE

IEEE COMMUNICATIONS SOCIETY

Conference Website
http://www.ieee-ctw.org/2012/

Shannon Award Call for Nominations

The purpose of the Claude E. Shannon Award is to honor consistent and profound contributions to the field of information theory. The selection is governed by Article V, Section 4.

An honorarium of $10,000 and a suitable memento are awarded to the Claude E. Shannon Award winners. Each Shannon Award winner is expected to present a Shannon Lecture at the IEEE International Symposium on Information Theory of the year of the award. In addition to the honorarium, the Information Theory Society will pay the winner's travel expenses.

The Shannon Lecturers in the years preceding the institution of the Shannon Lecturer Award (1973-1994) shall be considered to be Claude E. Shannon Award winners for the years their respective Shannon Lectures were delivered.

Nominations for the Claude E. Shannon Award can be made by anyone and are made by completing a nomination form (available online) and sending it and all supporting materials to the Society President by March 1st. The committee may consider all possible candidates, not only those for whom nominations have been received.

The 2012 C. E. Shannon Award Selection Committee, whose task is to decide whether to name a C. E. Shannon Award winner for 2013, will consist of the following members:

Muriel Medard (Chair)
Gerhard Kramer (ex officio)
Abbas El Gamal (ex officio)
Richard E. Blahut
Sergio Verdu
Raymond W. Yeung
Jacob Ziv

Aaron D. Wyner Distinguished Service Award Call for Nominations

The purpose of the Aaron D. Wyner Distinguished Service Award is to honor individuals who have shown outstanding leadership in—and provided long-standing exceptional service to—the Information Theory Community. The selection is governed by Article V, Section 9.

Nominations for the Wyner Distinguished Service Award can be made by anyone and are made by completing a nomination form (available online) and sending it and all supporting materials to the Society President by March 1st.

The individual or individuals making the nomination have the primary responsibility for justifying why the nominee should receive this award. The committee may consider all possible candidates, not only those for whom nominations have been received. Current officers and members of the Society Board of Governors are ineligible.

The prize shall be an ISIT or ITW participation fee waiver, a specially engraved plaque and a certificate, and shall be presented at the ISIT meeting held during the Summer following selection of the winner or at an appropriate IEEE IT society activity selected by the recipient.

The 2012 A. D. Wyner Award Selection Committee will consist of the following members:

Muriel Medard (Chair, ex officio)
Giuseppe Caire (ex officio)
Anthony Ephremides
Rolf Johannesson
H. Vincent Poor

Conference Calendar

DATE	CONFERENCE	LOCATION	WEB PAGE	DUE DATE
December 5–9, 2011	**2011 IEEE Global Communications Conference (GLOBECOM 2011)**	Houston, TX, USA	http://www.ieee-globecom.org	Passed
December 12–16, 2011	**15th Workshop on Quantum Information Processing (QIP2012)**	Montreal, Quebec, Canada	http://www.iro.umontreal.ca/~qip2012	Passed
February 5–10, 2012	**2012 Information Theory and Applications Workshop**	San Diego, CA, USA	http://ita.ucsd.edu/workshop.php	By Invitation
February 29–March 2, 2012	**2012 International Zurich Seminar on Communications**	Zurich, Switzerland	http://www.izs.etzh.ch/	Passed
March 21–23, 2012	**46th Annual Conference on Information Sciences and Systems (CISS 2012)**	Princeton, NJ, USA	http://ee-ciss.princeton.edu	January 6, 2012
March 25–30, 2012	**IEEE INFOCOM 2012**	Orlando, FL, USA	http://www.ieee-infocom.org	Passed
May 6–9, 2012	**2012 IEEE 75th Vehicular Technology Conference (VTC2012-Spring)**	Yokohama, Japan	http://www.ieeevtc.org/vtc2012spring	Passed
May 14–16, 2012	**2012 IEEE Communication Theory Workshop (CTW 2012)**	Ka'anapali, Maui, HI, USA	http://www.ieee.ctw.org/	March 1, 2012
May 14–18, 2012	**10th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt 2012)**	Paderborn, Germany	http://www.wi-opt.org/	January 6, 2012
June 10–15, 2012	**IEEE International Conference on Communications (ICC 2012)**	Ottawa, Canada	http://www.ieee-icc.org/	Passed
June 29–30, 2012	**International Symposium on Network Coding (NETCOD 2012)**	Cambridge, MA, USA	http://www.netcod2012.org/doku.php	February 24, 2012
July 1–6, 2012	**2012 IEEE International Symposium on Information Theory (ISIT 2012)**	Cambridge, MA, USA	http://isit12.org/	February 3, 2012
August 27–31, 2012	**7th International Symposium on Turbo Codes & Iterative Information Processing**	Gothenberg, Sweden	http://www.ee.kth.se/turbo-symposium-2012/	March 9, 2012
September 3–6, 2012	**2012 IEEE 76th Vehicular Technology Conference (VTC2012-Fall)**	Quebec City, Canada	http://www.ieeevtc.org/vtc2012fall/	February 2012
September 3–7, 2012	**2012 IEEE Information Theory Workshop (ITW 2012)**	Lausanne, Switzerland	http://itw2012.epfl.ch/	April 2, 2012
October 28–31, 2012	**2012 International Symposium on Information Theory and its Applications (ISITA 2012)**	Honolulu, HI, USA	http://www.isita.ieice.org/2012	March 28, 2012

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