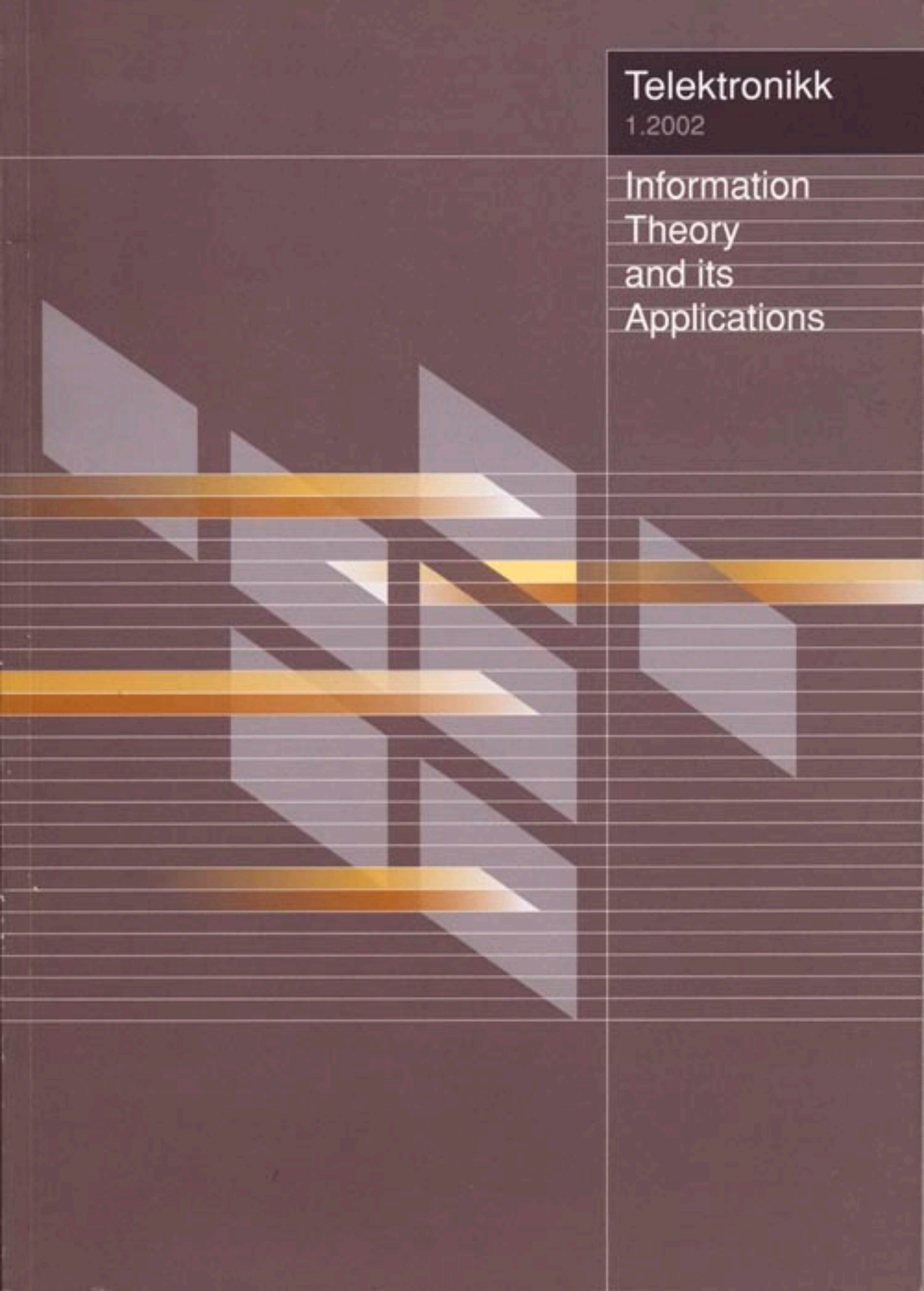


Teletronikk

1.2002

Information
Theory
and its
Applications

The cover features a dark brown background with a complex geometric pattern. This pattern consists of several overlapping, semi-transparent triangles in shades of grey and white, arranged in a way that creates a sense of depth and movement. Superimposed on this pattern are four prominent horizontal bars with a golden-yellow gradient, which appear to be layered over the other elements. The overall aesthetic is modern and technical, consistent with the journal's focus on information theory.

Contents

Teletronikk

Volume 98 No. 1 – 2002
ISSN 0085-7130

Editor:

Ola Espvik
Tel: (+47) 913 14 507
email: ola.espvik@telenor.com

Status section editor:

Per Hjalmar Lehne
Tel: (+47) 916 94 909
email: per-hjalmar.lehne@telenor.com

Editorial assistant:

Gunhild Luke
Tel: (+47) 415 14 125
email: gunhild.luke@telenor.com

Editorial office:

Telenor Communication AS
Telenor R&D
NO-1331 Fornebu
Norway
Tel: (+47) 67 89 00 00
email: teletronikk@telenor.com

Editorial board:

Berit Svendsen, CTO Telenor
Ole P. Håkonsen, Professor
Oddvar Hesjedal, Director
Bjørn Løken, Director

Graphic design:

Design Consult AS, Oslo

Layout and illustrations:

Gunhild Luke and Britt Kjus,
Telenor R&D

Prepress and printing:

Optimal as, Oslo

Circulation:

3,750

Feature: Information Theory and its Applications

1 Guest Editorial; *Geir E. Øien*

A Historical Perspective on Information Theory

3 Information Theory: The Foundation of Modern Communications;
Geir E. Øien

20 On Shannon and “Shannon’s Formula”; *Lars Lundheim*

30 Statistical Communication Theory 1948 – 1949; *Nic. Knudtzon*

Novel Developments on Channel Capacity and Spectral Efficiency

35 The True Channel Capacity of Pair Cables With Respect to Near End
Crosstalk; *Nils Holte*

47 Bounds on the Average Spectral Efficiency of Adaptive Coded Modulation;
Kjell J. Hole

53 Breaking the Barriers of Shannon’s Capacity; An Overview of MIMO
Wireless Systems; *David Gesbert and Jabran Akhtar*

Turbo Coding and Iterative Decoding: Theory and Applications

65 An Introduction to Turbo Codes and Iterative Decoding; *Øyvind Ytrehus*

78 Theory and Practice of Error Control Coding for Satellite and
Fixed Radio Systems; *Pål Orten and Bjarne Risløw*

Modulation, Coding and Beyond

92 A New Look at the Exact BER Evaluation of PAM, QUAM and
PSK Constellations; *Pavan K. Vitthaladevuni and Mohamed-Slim Alouini*

106 Performance Analysis of Adaptive Coded Modulation with Antenna Diversity
and Feedback Delay; *Kjell J. Hole, Henrik Holm and Geir E. Øien*

114 Shannon Mappings for Robust Communication; *Tor A. Ramstad*

Historical Papers

129 Information Theory; *Nic. Knudtzon*

139 Statistical Communication Theory; *Nic. Knudtzon*

145 Statistically Optimal Networks; *Nic. Knudtzon*

Special

153 Multiple Bottom Lines? Telenor’s Mobile Telephony Operations in
Bangladesh; *Arvind Singhal, Peer J. Svenkerud and Einar Flydal*

Status

163 Introduction; *Per Hjalmar Lehne*

164 UMTS Network Domain Security; *Geir M. Kjøien*

Guest Editorial

GEIR E. ØIEN



Geir E Øien

Front cover:

Information appears as a change in the detectable pattern

The artist Odd Andersen visualises a set of planes as areas for information to appear. Whatever kind of predictable pattern that already may exist on those planes has no interest. Only when that pattern is changed has information occurred. When part of the pattern of one plane is moved through a transmission channel to another plane and changes its pattern, this is seen as information received by the new plane.

The artist's generic message: Information has been produced and understood when the pattern of one plane has become changed and detected.

Ola Espvik, Editor in Chief

On February 24, 2001, *Claude Elwood Shannon* died at the age of 84. As the father of, and most important contributor to, the field of information theory, he ranks as one of the most brilliant and important of all 20th century scientists. In a tribute speech made by senator John D. Rockefeller in the US Congress after Shannon's death, his work was referred to as "The Magna Carta of the information age". Had there been awarded a Nobel prize within the information and communication sciences, no one would have been a more worthy candidate than Shannon.

Shannon's ideas, first presented to the world at large in his seminal 1948 paper "*A Mathematical theory of communication*" in Bell System Technical Journal, have been crucial in enabling the information and communication technological advances which have created today's information society. As with all true pioneers, Shannon's way of thinking about information and communication represented a true paradigm shift. For example, prior to the arrival of his seminal papers of the late 40s and 50s, there simply had been no satisfactory way of modelling and analyzing the process of information generation, transfer, and reception from a transmitter to a receiver over a noisy communication channel – which actually is a generic description of how all practical communication systems work.

With Shannon's introduction of a generic communication system model, his view of information as a probabilistic entity (sidestepping its actual semantic meaning), the insight that the process of information transmission is fundamentally stochastic in nature, and his invention of precise mathematical tools to give a complete performance analysis of his model, the door was suddenly opened to a much more fundamental understanding of the possibilities and limitations of communication systems. In the words of another notable information theorist (and former colleague of Shannon), David Slepian, "*Probably no single work in this century has more profoundly altered man's understanding of communication than C.E. Shannon's article, "A mathematical theory of communication", first published in 1948. The ideas in Shannon's paper were soon picked up by communication engineers and mathematicians around the world. They were elaborated upon, extended, and complemented with new related ideas. The subject thrived and grew to become a well-rounded and exciting chapter in the annals of science.*"

In a world where predictions for the future performance of telecommunication systems sometimes seem to be made more out of marketing concerns than out of a scientifically sound judgement, information theory still has a lot to teach us – insights that are sometimes sobering, but may also be encouraging. The optimistic vision suggested by some, that "any telecommunication service can be made available anywhere, anytime, and to anyone" in the future, can fairly easily be shown to have no roots in reality. One example may illustrate this: The claims originally made regarding the available rates and coverage for the upcoming Universal Mobile Telecommunications System (UMTS) so far seem a bit over-optimistic ...

However, in some cases information theory can also be used to uncover a performance potential beyond what was previously thought possible, and aid in the design of systems realizing this potential. As an important example of the applicability of Shannon's results, his theory showed us how to design more efficient communication and storage systems by demonstrating the enormous gains achievable by coding, and by providing the intuition for the correct design of coding systems. The sophisticated coding schemes used in systems as diverse as deep-space communication systems, and home compact disk audio systems, owe their success to the insights provided by Shannon's theory.

Shannon published many more important and influential works in a variety of disciplines, including Boolean algebra and cryptography. His work has had an influence on such diverse fields as linguistics, phonetics, psychology, gambling theory, stock trading, artificial intelligence, and digital circuit design. It also has strong links to disciplines such as thermodynamics and biochemistry.

Shannon was also known for his playfulness and eclectic interests, which led to famous stunts such as juggling while riding a unicycle down the halls of Bell Labs. He designed and built chess-playing, maze-solving, juggling, and mind-reading machines. These activities bear proof to Shannon's claim that his motivation was always curiosity more than usefulness. In an age where basic research motivated purely by scientific curiosity sometimes seem to be at the losing end as far as public interest and funding is concerned, this is a statement worth remembering.

ing. The success of information theory clearly shows how one person's curiosity can translate into very useful results.

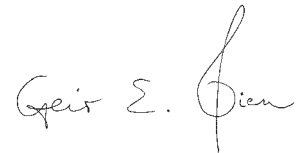
In Norway, Shannon's memory and the achievements of information theory were recently honored with a *Claude E. Shannon In Memoriam Seminar*, arranged by Telenor R&D at Kjeller on the August 9, 2001. The seminar drew over 100 participants from Norwegian industry and academia. They came to listen to technical talks by some of Norway's foremost experts within the fields of information and communication theory, as well as to some historically and philosophically flavored talks. The seminar was highlighted by a unique reminiscence by former Telenor R&D director Nic. Knudtzon, the only Norwegian to have met Shannon at the pioneering time when information theory was actually born. This special issue of *Teletronikk* was inspired to a great extent by the success of this seminar, and you will find papers by many of the same scientists here – including Dr. Knudtzon's wonderful reminiscence, as well as three 1950 papers of his which rank as the very first presentations of information theory made in Norway.

When putting together this issue, I set myself one main goal and four sub-goals. The main goal was to collect a high-quality collection of papers which could serve to dispel the often-heard view that "information theory is not constructive, but is only concerned with theoretical limits which can never be reached in practice". It is my firm view that information theory can be *very* constructive, in the sense that it has a lot to say about what one should do and what one should *not* do when designing communication systems.

In fact, these are particularly exciting times for information theory, because we now finally have available powerful techniques for actually approaching the performance limits predicted by Shannon. One important example is the class of error control codes called "turbo codes". Why do these codes work so well? Because their construction turns out to be based to a great extent upon the "nonconstructive" proof techniques used by Shannon in his analysis!

Regarding my four subgoals; first, the issue should place information theory in a historical context, while at the same time introducing its basic principles and discussing their importance. The papers collected under the heading "*A historic perspective on information theory*" serve this purpose.

Secondly, I wanted to show that information theory is still very much an active research field of practical importance, and that the Norwegian research community is currently making some important contributions to this research. Thirdly, some of the most important applications of information theory should be highlighted. Finally, the most promising and exciting current developments in communication systems design should be included. I do believe that the final result reflects these goals successfully, and that the standard of the papers within is very high. I hope the reader will find them as inspiring, insightful, and useful as I do. For me, putting together this issue has truly been a labour of love!



Information Theory: The Foundation of Modern Communications

GEIR E. ØIEN



Geir E Øien (36) received his MSc and PhD degrees from the Norwegian University of Science and Technology (NTNU) in 1989 and 1993, respectively. From 1994 until 1996 he was with Stavanger University College as associate professor. Since 1996 he has been with NTNU, since 2001 as full professor of information theory. Prof. Øien is a member of IEEE and the Norwegian Signal Processing Society. He is author/co-author of more than 40 research papers. His current research interests are in information theory, communication theory, and signal processing, with emphasis on analysis and design of wireless communication systems.

oien@tele.ntnu.no

This paper gives a brief tutorial introduction to *information theory*, the fundamental concepts and theorems of which are at the very heart of modern communications and information technology. The results of information theory tell us:

- how to quantify the information content in a set of data;
- how to model and analyze a wide range of communication channels and their capacity for transmitting information;
- conditions under which error-free representation and transmission of information is possible, and when it is strictly impossible;
- conditions for the design of good ways of (codes for) representing information so as to achieve data compaction and compression, and channel error robustness;
- which minimal quality reduction we may expect for a given transmission rate, a given information source, and a given communication channel;
- how we may split our communication systems into subsystems, in order to simplify design without the loss of theoretical performance.

We will present the basic concepts and most famous theorems of information theory and explain their usefulness in the design and analysis of communications and information storage systems.

1 Introduction

Modern information theory was basically invented by Claude E. Shannon in a series of classic papers [1], [2], [3], and has since been extended and refined by a large number of researchers from all over the world. To some, the field of information theory might seem overtly concerned with mathematical and statistical rigor, and with an abundance of long, technically involved mathematical proofs. However, the advantages of this rigor are many. Starting out with mostly simple, but stringently defined, mathematical assumptions – typically corresponding to practical/physical constraints under which real-world communication systems are to be designed – one may derive fundamental limits on the performance of such systems, as well as making statements about the conditions under which a given performance can be attained. In many situations we are also able to obtain useful guidelines for how the practical design should (not to mention should *not*) be done in order to approach the theoretical performance bounds.

In the following, we shall define and explain some of the most basic concepts and results introduced by Shannon, and try to expand on how they affect the design of communication and information storage systems. Due to space limitations and to ease the reading of this paper, we shall not provide proofs for any of the results presented. Most of the notation and language is based on that of Blahut [4] and the more recent book by Cover and Thomas [5]. Another important text on information theory is Berger's classic on rate distortion theory [6].

2 Shannon's Generic Model of a Communication System

Central to the development of information theory is the notion of a *generic communication system model* which can be used as a unifying framework suitable for describing a wide range of real-world systems. In the generic communication system model proposed by Shannon, information is transmitted from an *information source* to a *user*, by means of a *transmitter*, a

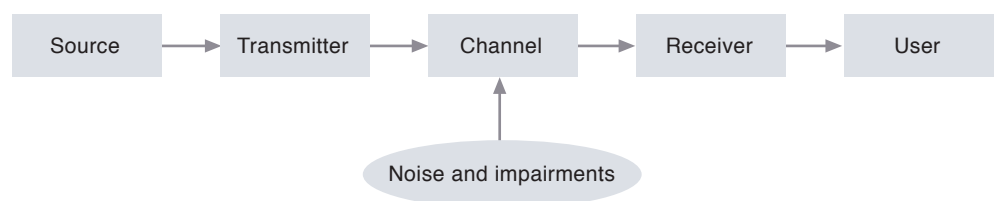


Figure 1 Generic communication system model

communication channel, and a *receiver*. This is depicted in Figure 1. Examples of possible information sources in this context are: human speakers, video cameras, musical instruments, microphones, loudspeakers, and computer keyboards.

The transmitter and receiver perform information *coding*, which means processing the messages generated by the information source in order to

- represent (*encode*) the messages in a suitable way during transmission over the channel, and
- regenerate (*decode*) the messages at the receiver end, with as little deviation from what was originally transmitted as is required for the service under discussion.

Note that the term “transmission” here is intended to cover transmission both in space (between two different locations) and in time (i.e. *storage of data* on an imperfect medium). Examples of physical communication channels thus range from wireless channels such as satellite links, mobile radio channels, and broadcast channels, via wired links such as optical fibres and copper transmission lines, to magnetic storage media and CDs.

The impairments to the transmitted messages may differ a lot depending on the type of channel, but may include

- thermal noise and atmospheric noise (additive noise);
- interference from other sources and system users;
- reflections and scattering of transmitted radio wave power during propagation through the terrain;
- signal attenuation due to path loss in radio channels or transmission line resistance;
- intersymbol-interference due to a lack of available bandwidth;
- Doppler shifts due to relative movements between receiver and transmitter;
- nonlinear effects, e.g. due to nonlinear power amplifier characteristics;
- stains or scratches on a CD-ROM.

Whatever the channel, the aim of information theory is to model these impairments in a quantitative way, mainly by statistical models. The modelling is then used to deduce performance

limits, and to devise methods for efficient transmission over the channel – that is, *coding* algorithms.

The ultimate goal of the coding is to exploit the communication channel as well as possible. This is to say that we want to spend as little as possible of the limited physical resources we have available – time, bandwidth, transmit power, or disc space – on the transmission or storage of information, in order to maximize the number of users, systems, or services that are able to share these resources. At the same time we need to ensure that the *quality* of the information retrieved at the receiver end is satisfactory.

By “quality” we usually mean something like “degree of similarity to the transmitted message”. The appropriate measure of, and typical demands on, quality is dependent on the type of information transmitted and on the application or service: For data communications, the criterion might be that the *average information bit error probability* (bit error rate – BER) should be less than, say, 10^{-9} – whereas for speech and video communications, the aural or visual quality perceived by human ears or eyes is the most important thing, and a much higher BER can usually be accepted.

The accepted perceptual quality range also varies with the application, and is different e.g. for mobile telephony (low-to-medium quality application) and high fidelity audio (very high quality application). Real-time applications such as two-way speech communication also place constraints on average delay, buffering, probability of no transmission (“outage”), etc.

3 Information – What is it?

One of the most basic questions answered by information theory is “What is information – in a quantitative sense?” When discussing this issue it will be useful to distinguish between two different perspectives:

- How to quantify the information content in the data produced by a source (as discussed in Subsection 3.2).
- How to quantify the information content transmitted from a source to a user by means of a channel (as discussed in Section 5).

3.1 Information as a Measure of Unpredictability

Regardless of which of the two above perspectives is used, an intuitively appealing line of thinking about information content is as follows: An event has a low information content if its “future” can be forecast with a high degree of accuracy, based on knowledge of its “past”. In

this case an observer's *uncertainty* about the future of the event is low, which means he will not receive much new information by continued observation of the event.

Thus, an event's information content is intimately linked to the *a priori degree of randomness and uncertainty* associated with the event: The more predictable an event is, the less knowledge we need to describe or predict it; thus the less information we receive by observing it.

Conversely, if an event is highly unpredictable, a detailed observation of its actual development is needed in order to describe it, which means its information content is high. This holds *regardless of the "physical" content of the event in question* – sometimes referred to as the *semantic meaning* of the information.

The above can be restated as saying that the actual semantic meaning in a message is of no consequence for the amount of information carried by the message. All that matters is the degree of predictability; i.e. its statistical properties.

Note, however, that the predictability of an event is usually strongly linked to our knowledge of physical or statistical properties and models of the event in question. The same event may appear unpredictable to one observer and highly predictable to another, the difference being that only the latter observer has a priori knowledge of the underlying model describing the event.

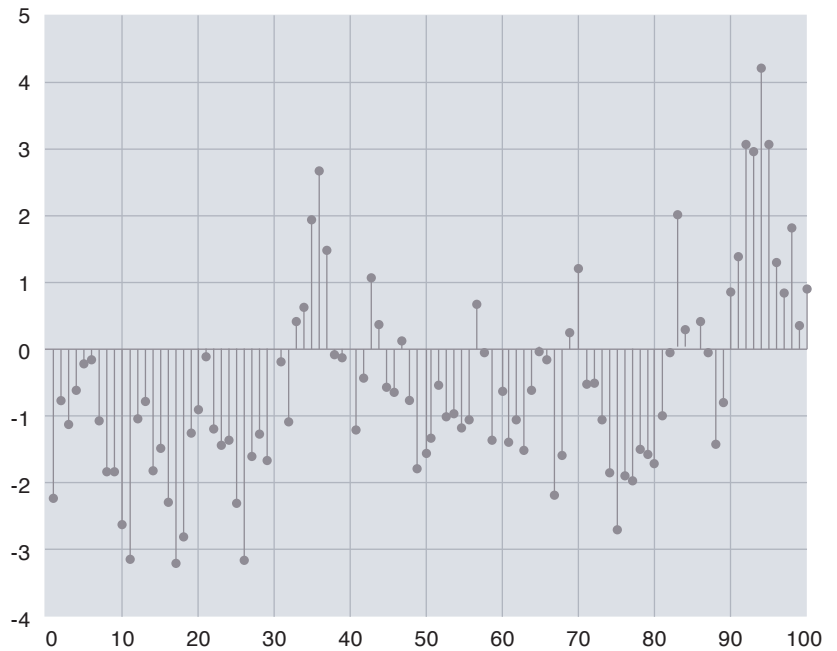


Figure 2 Example of time-discrete source output

An example of this is found in language modelling, applied e.g. in speech recognition, where an observer will need knowledge of the structure (usually in the sense of *Markov models* described by transition probabilities) and vocabulary of a language if he is to attempt predicting future words in a sentence, based on what has already been said.

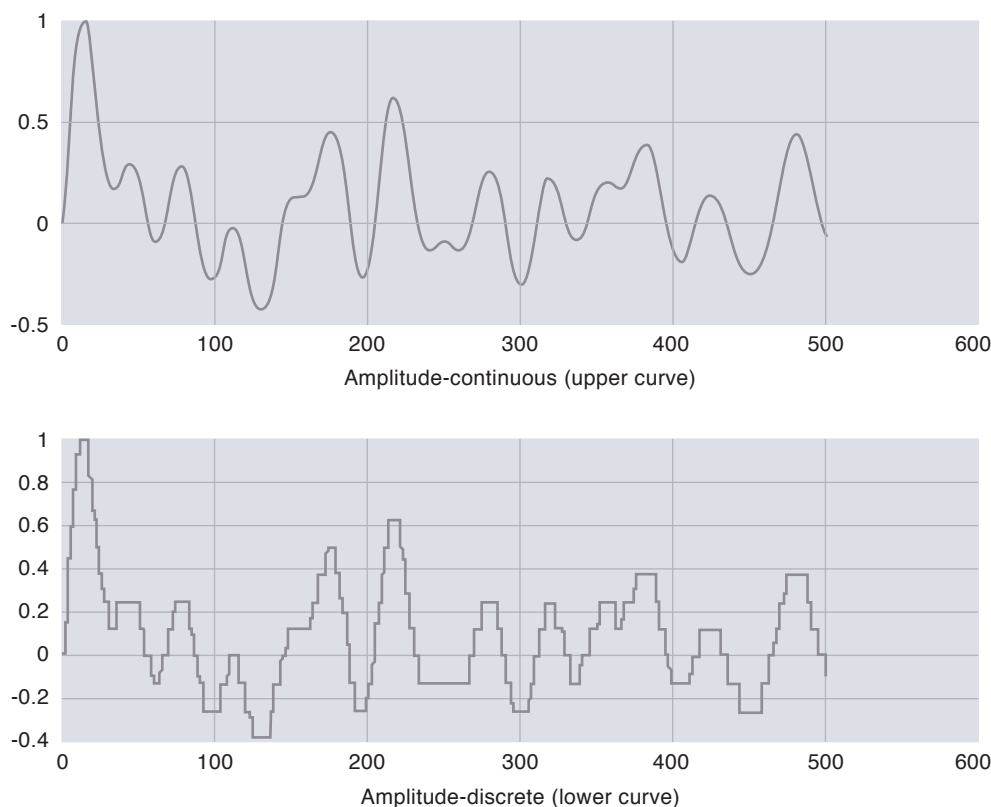


Figure 3 Examples of continuous-time source outputs which are respectively amplitude-continuous (upper curve) and amplitude-discrete (lower curve)

In a communication context, at least from the second of the two above perspectives, the “events” under study are usually information messages passed from a source to a user. The degree of uncertainty on the receiver side, and the a priori predictability of the messages, will then depend on the statistical properties of the information source under study, the impairments introduced by the physical communication channel, and the way the transmitter and receiver are designed.

3.2 The Information Content in a Source

Armed with these basic insights, we now address the following question: “What is the information content in a given source that outputs messages which we are interested in storing, compressing, or transmitting?”

Initially, to make the explanations and notation simple while still illuminating the general theory, we will consider an information source that has been *digitized*, such that its output is discrete both in time (sampled) and amplitude (quantized). Figure 2 shows outputs from a discrete-time source, while Figure 3 shows examples of continuous-time source outputs that are amplitude continuous and discrete, respectively. When both time and amplitude have been discretized we say that the source output consists of *discrete-valued samples*, and the source is referred to as *discrete*.

We denote the set of J possible sample values (also called *representation levels*, or more generally *source symbols*) by $S = \{x_1, \dots, x_J\}$. This set is called the *source alphabet*. We assume that

the probability of the source output x_j is known (e.g. through histogram estimation based on training data) to be p_j .

3.2.1 Source Entropy

Let us now introduce a notion which will turn out to be intimately linked to source information content: The *entropy* of a source S as described above is defined as

$$H(S) = - \sum_{j=1}^J p_j \log_2 p_j, \quad (1)$$

and is measured in *information bits per source symbol*. Note that the base of the logarithm in the general definition of entropy is arbitrary, but for the entropy to be measured in units of bits per sample, the binary logarithm must be used.

An illustration of the entropy for the special case of a *binary* source ($J = 2$) is shown as a function of symbol probability p (the two symbols in this case have probabilities p and $1 - p$, since their probabilities must sum to 1) in Figure 4. Note that the entropy reaches its maximum – of 1 information bit per symbol – when the two possible outcomes are equiprobable, i.e. $p = 0.5$.

3.2.2 Entropy as a Measure of Information Content

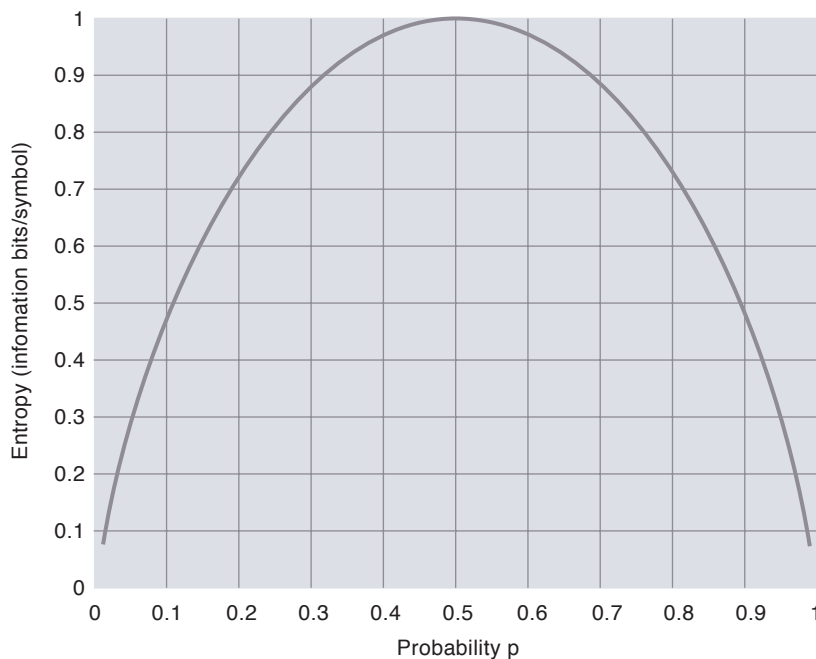
A commonly used information theoretic statement is that “the source entropy (as defined above) is a measure of the average amount of information per symbol produced by the source.” What does this mean? *Shannon’s first coding theorem* (the noiseless source coding theorem) addresses this question in a precise way [5].

In the simplest version of this theorem, it is assumed that the source in question is *memory-less*, i.e. that the different source symbols output by the source are *statistically independent* of each other. Also, for simplicity, we shall assume that we want to represent each source symbol by a distinct *binary word*, for storage or transmission on a channel where only binary symbols are admitted (such as is the case e.g. on a CD).

We denote by l_j the number of bits in a *code-word* which will be used to represent the source symbol x_j . The set of codewords for all j make up the *source code*. Reasonable demands on this code are that it should be able to encode any symbol string output by the source without errors, and that it should be *uniquely decodable*, i.e. an arbitrary encoded binary string should be possible to *decode* without errors into a unique sequence of source symbols.

In this situation $H(S)$ turns out to be the *lower limit* for the *average number of bits per symbol*

Figure 4 The entropy of a binary source



(average codeword length), \bar{l} , we may use to represent the source output without errors or ambiguities:

$$\bar{l} = \sum_{j=1}^J p_j l_j \geq H(S). \quad (2)$$

In other words, it is not possible to construct a code with a lower average codeword length than $H(S)$ bits per codeword. Conversely, it is theoretically possible to construct codes whose average length comes arbitrarily close to this limit (from above). This is the essence of the noiseless coding theorem.

Thus the notion of entropy as a measure of the source information content is rather obvious if viewed from a data storage viewpoint. Note that the entropy is maximal ($\log_2 J$ bits/symbol) if all source symbols are *equiprobable*, i.e. maximally unpredictable. It is minimal (0 bits/symbol) for a *deterministic source*¹⁾, whose future output “tells us nothing new” and is thus totally redundant.

4 Source Code Construction

Of course, one thing is to know the attainable limit of code efficiency, quite another is to actually find a code approaching this limit. Happily, information theory also provides necessary and sufficient conditions for how to *construct* the codewords of an optimal code, which actually approaches the above lower codeword length limit as closely as possible [4].

We will here still consider only the (practically quite interesting) case of a *binary, variable-length, and prefix-free* code – i.e. a code whose codewords are binary strings, where different codewords might have different lengths, and in which no codeword is a *prefix* of another. The practical interest of this last property is to make the code *instantaneously decodable*, i.e. each codeword can be immediately decoded without any need for “look-ahead” to other codewords.

A necessary condition for a variable-length prefix-free code to be *uniquely* decodable is that all of its codewords can be put on a *code tree*. An example of such a tree is shown, for the example of a source alphabet with four symbols, in Figure 5. The way to obtain a codeword from the tree is to start at the top node and write down the binary-valued labels as the tree is traversed down to one of the end nodes. Each resulting sequence of labels then constitutes one codeword – one for each end node. It is easily seen that the set of codewords thus constructed has the prefix-free property.

4.1 Huffman Coding

As stated previously, there exists a constructive algorithm which may be used to build an optimal code according to the above principles. The resulting code is the well-known *Huffman code* [7].

The Huffman algorithm is best illustrated by means of an example: Consider a source with 5 different possible symbol outcomes x_1, \dots, x_5 , with corresponding probabilities given by

$$\begin{aligned} p_1 &= 0.30 \\ p_2 &= 0.24 \\ p_3 &= 0.20 \\ p_4 &= 0.15 \\ p_5 &= 0.11 \end{aligned} \quad (3)$$

The Huffman code construction for this case is illustrated in Figure 6. As seen from the figure, the symbol probabilities are initially ordered from the largest to the smallest. The two least probable symbols (to begin with, symbols 4 and 5 in this case) are assigned a distinct binary code symbol each – 0 or 1 – to be able to distinguish between them. Then their probabilities are added, to form the probability of a “quasi-symbol” whose outcome is defined as “one of the two outcomes whose probabilities have just been added”. The code symbols just assigned can be used to distinguish between the two possible quasisymbol outcomes.

Then the remaining probabilities are re-ordered, with the quasi-symbol probability being inserted at its proper place according to its size. The above procedure is thereafter repeated for the remaining symbols/quasi-symbols. Repeating the procedure several times, in the end there

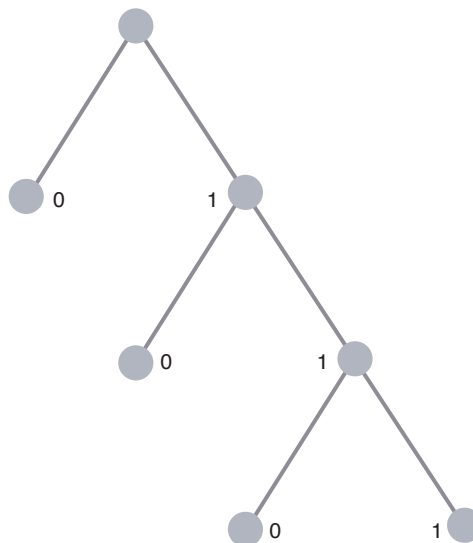


Figure 5 Code tree for the binary prefix-free code $\{0, 10, 110, 111\}$

¹⁾ This is the case e.g. if $p_1 = 1$ and $p_2 = \dots = p_J = 0$.

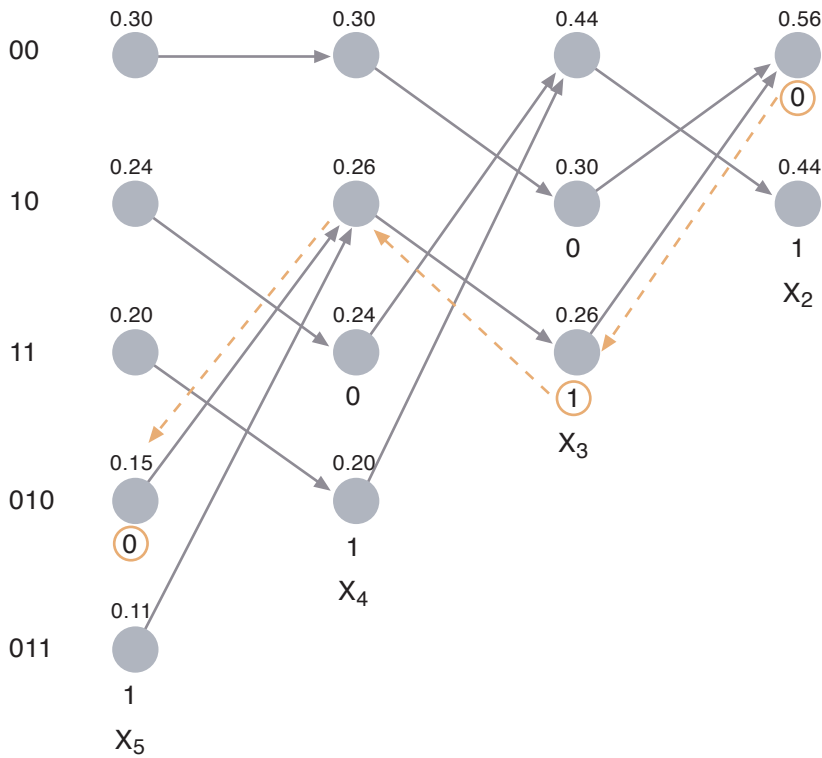


Figure 6 Construction of Huffman code for a 5-symbol source. Two arrows pointing to a single node denotes the adding of two probabilities to form a “quasisymbol”. The backtracing to find the codeword 010 for the symbol X_4 is shown as dashed arrows, with corresponding binary labels marked by circles

will be only two symbols/quasi-symbols left (such as is the case to the right in Figure 6), of which one is assigned a 1 and the other a 0. The codewords which are used to distinguish between all original symbol outcomes can now finally be deduced by tracing the paths from the last two symbols back to each original symbol, while at the same time noting the binary labels given at each forward repetition of the procedure. In this example such backtracing gives the codes

$$\begin{aligned}
 C_2 &= \{0, 1\} \text{ for the two-symbol quasi-source,} \\
 C_3 &= \{1, 00, 01\} \text{ for the three-symbol quasi-source,} \\
 C_4 &= \{00, 01, 10, 11\} \text{ for the four-symbol quasi-source,} \\
 C_5 &= \{00, 10, 11, 010, 011\} \text{ for the original five-symbol source.}
 \end{aligned}
 \tag{4}$$

C_5 is the Huffman code for the source.

4.2 Source Extension

In general, to obtain a code with an average codeword length coming arbitrarily close to the entropy, one must combine Huffman coding with *source extension*, a process in which *vectors* of source symbols are treated as single symbols from an “extended” source. Using vectors of length n , the extended source has J^n vector symbols. As an example, a ternary (3-dimensional) extension of a binary source with alphabet $\{0, 1\}$ will yield a source with binary vector outcomes in the alphabet $\{000, 001, 010, 011, 100, 101, 110, 111\}$.

If the source is assumed memoryless, each vector is associated with a probability equal to the *product* of the probabilities of its individual elements. A Huffman code may now be constructed according to this new, extended set of probabilities. The larger n is used, the closer to the entropy we are able to come; in fact it is easy to show that the following bounds on the average codeword length hold for a Huffman code when n -fold source extension is used [5]:

$$H(S) \leq \bar{l} \leq H(S) + \frac{1}{n} \text{ [bits per symbol]} \tag{5}$$

The above result is dependent on the fact that we know the symbol probabilities. Thus a limitation of the Huffman code is that the source’s probability distribution has to be known or estimated a priori in order for a code to be constructed. If this distribution is estimated incorrectly, the code will be less efficient – the average codeword length will be increased by an amount which can also be quantified exactly using information theoretic notions [5].

However, there also exist source code constructions which are able to asymptotically (as the number of symbols to be coded goes to infinity) reach the entropy bound even without any a priori knowledge of the source’s statistical properties. Such codes are called *universal codes*. The *Lempel-Ziv* algorithm [5] for compaction of large data files is by far the most used universal code. Note that universal codes do not necessarily provide compression for short symbol sequences.

4.3 Sources with Memory

Most natural information sources are *not* memoryless; rather, the samples are correlated with each other. Indeed, this is what makes it possible for us as observers to make sense of the source output. A good example is again that of human languages: It is precisely the fact that there exist well-defined grammatical rules and structure, which lead to interword and inter-sentence correlation (i.e. memory, or statistical dependencies) which makes it possible to convey and understand information by means of a written text or a spoken message. How does the above theory apply to sources with memory?

4.3.1 Entropy Rate

The answer can loosely be said to be as follows: By applying source extension with sufficiently large vector lengths, *any* random source is essentially made into a memoryless extended source. There may be correlation or statistical dependency between individual elements within each vector, but not between two different vectors – or “extended symbols” – if these are long enough. To make this argument stringent in the

general case, one needs to let the vector length n go to infinity, to account for correlation over arbitrary lags. The minimum average number of bits per original source symbol is now equal to the *entropy rate* of the source,

$$\mathcal{H}(S) = \lim_{n \rightarrow \infty} H(S^{(n)}), \quad (6)$$

where $H(S^{(n)})$ is the entropy of the source S extended to n -dimensional “supersymbols”. The entropy rate thus replaces the first order entropy as the natural measure of information content in a source with memory.

4.3.2 Markov Sources

One common way of modelling memory in sources is to use *Markov models*. This is an accepted model e.g. for natural speech production. A (discrete-time, discrete-amplitude) Markov model of order M has the property that its memory stretches back only M time instances. More precisely, the probability of a symbol outcome at time m is dependent on the outcomes at times $m-1, \dots, m-M$, but not further back in time. Each possible set of the M previous outcomes constitutes a model *state*. The number of possible states is less than or equal to J^M , where J is the number of possible outcomes in the source alphabet. Statistically, the model is completely characterized by *transition probabilities* which together quantify the probability of each state being followed by each of the other states.

A Markov model is usually visualized by means of a *state diagram*. A state diagram of a simple 1st order Markov model with $J=3$ source symbols $\{a_1, a_2, a_3\}$ is depicted in Figure 7. Here,

$P_{a_i|a_j}$ is by definition the probability of the transition from a_j to a_i . If there is no arrow between two of the states, it means that that particular transition is impossible to make.

For Markov sources, there is a simple expression for the entropy rate. It is simply the expected value – with respect to the state distribution – of the entropy *conditioned on being in a given state*. Denoting the n th state by s_n we obtain:

$$\mathcal{H}(S) = \sum_{n=1}^{J^M} P(s_n) H(X|s_n) \quad (7)$$

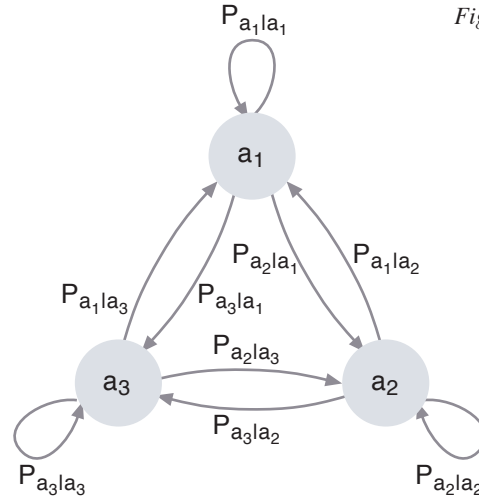


Figure 7 State diagram of a 1st order, 3-state Markov source

where $P(s_n) = \sum_{l=1}^{J^M} P(s_l) P(s_n|s_l)$, with

$\sum_{l=1}^{J^M} P(s_l) = 1$. This set of equations can be used to solve for the state probabilities $P(s_l)$, $l=1, \dots, J^M$, when the transition probabilities $P(s_n|s_l)$ are given. The conditional entropy $H(X|s_n)$ is given by the usual entropy formula, only with state-conditional symbol probabilities replacing unconditional symbol probabilities:

$$H(X|s_n) = - \sum_{j=1}^J P(x_j|s_n) \log_2 P(x_j|s_n). \quad (8)$$

Equation (7) can now be used to find the ultimate limit of error-free compressibility for a Markov source, and a set of *state-conditional Huffman codes* can be used to provide efficient source compaction. That is to say, for *each state* we can construct an optimal Huffman code as before, but according to the *conditional* symbol probabilities corresponding to the state. The encoder will then switch between the different Huffman codes according to the state of the source.

4.4 Continuous-Amplitude Sources

Most “natural” sources not only have memory, but are *analogue* in nature, and hence produce data that are continuous both in time (or space) and amplitude. Examples are acoustic and electromagnetic waves as encountered e.g. in audio technology and in radio communications. *Sampling* and *quantization* operations convert such sources into time and amplitude-discrete ones as discussed in previous sections. Of these operations, sampling can preserve all information as long as the source is band limited and the sampling theorem criterion, i.e. sampling at a rate of at least twice the bandwidth of the source, is fulfilled.

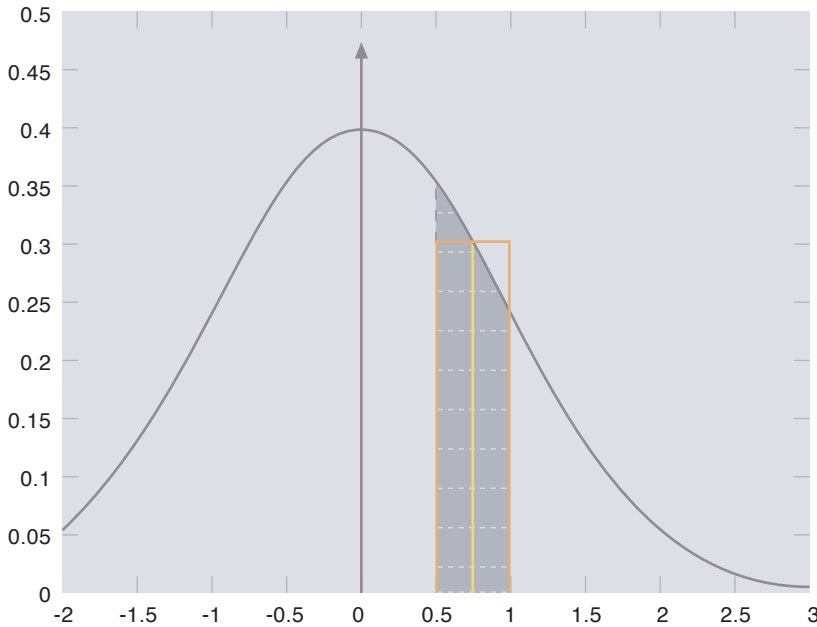


Figure 8 A Gaussian pdf with expected value 0, variance 1, with $P(0.5 \leq x \leq 1)$ shown coloured, together with the approximation (quantization) $p_X(0.75) \cdot 0.5$ (i.e. $\Delta = 0.5$)

Quantization, however, always reduces information, as measured through the source entropy. In fact, while the entropy of a discrete source is always finite, every nontrivial continuous-amplitude source has infinite entropy. This merely reflects the fact that with a continuous probability density, every fixed amplitude value has probability zero, so it is impossible to predict (or even measure) with infinite resolution which values forthcoming samples will have.

The way to see that this is true mathematically is to model the continuous amplitude distribution as the result of a uniform quantization, with a quantization interval Δ that goes to zero. This is illustrated for a Gaussian source in Figure 8.

The entropy may then be approximated by a sum over the quantized probability distribution. This sum converges to an integral over a source probability density function (pdf) $f_X(x)$ as Δ goes to zero. The expression for the amplitude-continuous source entropy then becomes

$$H(S) = \int_{-\infty}^{\infty} f_X(x) \log_2 f_X(x) dx - \lim_{\Delta x \rightarrow 0} \log_2 \Delta x \quad (9)$$

The last term is an infinitely large positive constant, which is common to all analogue sources. This implies that the absolute entropy of an analogue source will be infinite.

The first integral term in the above equation, however, is finite and well defined, and merits its own name and notation: It is the differential entropy, $h(S)$, of the source. The main reason for the importance of this term, which in itself has no fundamental physical interpretation²⁾, is the fact that several important information theoretic concepts, such as channel capacity, are defined in terms of differences between entropies.

In the continuous case, the difference between two (infinitely large) source entropies is always equal to the difference between the corresponding (finite) differential entropies, whose values can be computed. This situation arises because the term $-\lim_{\Delta x \rightarrow 0} \log_2 \Delta x$ is the same for all continuous sources, and thus is cancelled out whenever a difference is computed.

5 Information Transmission over a Channel

We now consider the second of the two perspectives previously introduced regarding information content: That of quantifying the amount of information transferred to a user from a source by means of a channel. In doing so, we need to introduce another information theoretic notion, that of mutual information. This is a function which can be said to measure the amount of useful source information received by a user. Here, it is important to emphasize the usability of information received, as the total “information content” received can stem from many sources, including noise and interference. This type of “information” is not only superfluous; as we shall see it actually reduces the useful information.

Mutual information is one of the most fundamental information theoretic concepts. We will show later that it can be used to describe not only the capacity properties of communication channels, but also the compression possibilities for a given source when a certain amount of error (distortion) is accepted in the source representation.

5.1 The Mutual Information Function

Initially consider two given discrete, possibly statistically dependent random variables, $X \in \{x_1, \dots, x_J\}$ and $Y \in \{y_1, \dots, y_K\}$. We can think of X as the input to, and Y as the output from a communication channel; i.e. Y is a “noisy” version of X . X and Y have probability distributions $\mathbf{p} = [p_1, \dots, p_J]^T$ and $\mathbf{q} = [q_1, \dots, q_K]^T$ respec-

²⁾ In some cases, though, it can be seen as a measure of the “relative randomness” of a given source compared to another: For example, for two different Gaussian sources, the one with the largest variance (and hence the largest amplitude variation from sample to sample) will have the largest differential entropy. However, it is not possible to successfully generalize this interpretation to two sources with different functional forms on their probability density.

tively. The *conditional* probability distribution of x_j , given y_k , is denoted $P_{j|k}$. $P_{j|k}$ will then describe the probability of x_j being transmitted if y_k is observed at the receiver.

The *mutual information between X and Y* is now defined as

$$I(X;Y) = H(X) - H(X|Y), \quad (10)$$

where $H(X)$ is the entropy of the source that outputs X , and, by definition,

$$H(X | Y) = \sum_{k=0}^{K-1} q_k \cdot H(X | y_k) = - \sum_{j=0}^{J-1} \sum_{k=0}^{K-1} q_k P_{j|k} \log_2 P_{j|k} \quad (11)$$

$H(X|Y)$ should be read “the entropy of X when Y is observed”. It is a *conditional* entropy, which means that it is based on the a priori knowledge of some information – in this case Y . It is simply a measure of the average information content (uncertainty) which is left in X when Y is known.

Then, the mutual information $I(X;Y)$ can be thought of as *the information Y gives about X* – the average reduction in the observer’s uncertainty about X which is brought about by observing Y . If X is the input to, and Y is the output from, a given communication channel, $H(X|Y)$ can then be thought of as information “lost” by the channel.

For an ideal (i.e. noiseless) channel Y would of course be equal to X , in which case it is easy to show from the definition that $H(X|Y)$ is zero. Hence $I(X;Y) = H(X)$, as of course is intuitively correct in this case: All transmitted information has been received; X has been fully determined by observing Y .

Correspondingly, if X and Y are *independent* variables, Y carries *no* information about X , and knowing Y cannot reduce our uncertainty about X . In a communication context this would correspond to the received signal being completely buried in noise. Hence it is intuitive that $H(X|Y) = H(X)$ – all information is lost during transmission. Thus the intuitive result would be $I(X;Y) = 0$ in this case. This also turns out to be the case

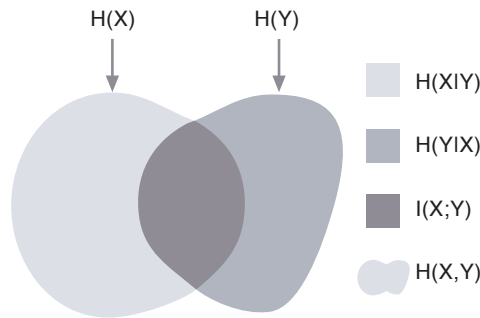


Figure 9 Visualization of the relationship between entropies, conditional entropies, mutual information and joint entropy for two variables X and Y

if the above formulae are applied. For most practical channels, $0 < I(X;Y) < H(X)$.

It is interesting to note that the mutual information is a *symmetric* function, i.e. $I(X;Y) = I(Y;X) = H(Y) - H(Y|X)$. In some cases, like when computing the channel capacity of a channel with input X and output Y , this might actually be a more useful form of the function. Physically, it means that the transmitter’s degree of uncertainty regarding what will be received, is the same as the receiver’s uncertainty regarding what was sent.

The above theory can be beautifully and simply illustrated by means of a Venn diagram, as shown in Figure 9. The *joint entropy* $H(X;Y)$ referred to in this figure is the total average uncertainty an observer will encounter regarding the *simultaneous* outcomes of X and Y .

5.2 Mutual Information in the Continuous Case

Mutual information is an example of a concept defined in terms of a difference between two entropies. If the variables X and Y are continuous random variables (with infinite absolute entropies), this difference reduces to a difference between *differential entropies*. The formula becomes

$$I(X;Y) = h(X) - h(X|Y) = - \int \int f_{XY}(x,y) \log_2 f_{X|Y}(x) dx dy$$

where $f_{XY}(x,y)$ is the joint pdf of X and Y . Both integrals are taken from $-\infty$ to ∞ . Through the use of *Bayes’ rule* [4] we may manipulate this

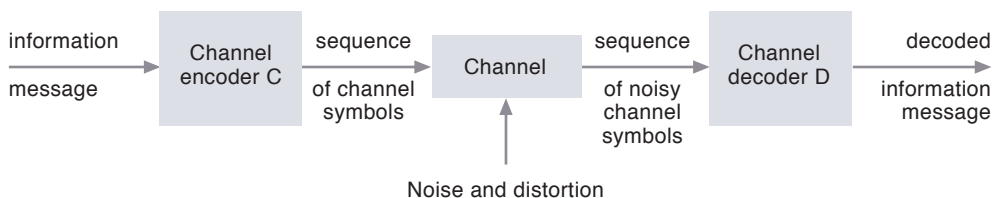
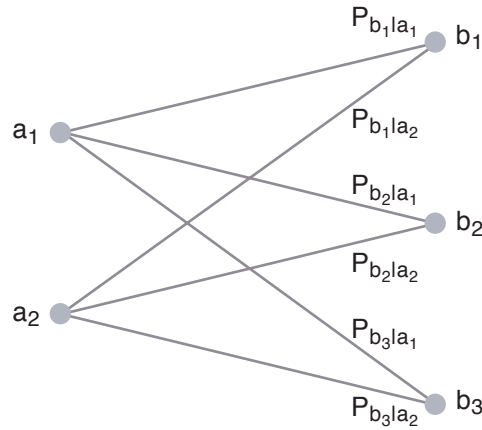


Figure 10 Channel encoding and decoding as mappings

Figure 11 Input-output transition diagram for discrete memoryless channel ($J=2, K=3$)



formula in various ways to incorporate the probability densities that are most practical to use in a given situation. We shall now see how the mutual information function can be used to find the *capacity* of a general communication channel.

6 Channel Capacity and Coding

As discussed in the introduction, most of the physical channels we transmit information over are subject to noise and distortion of various kinds, resulting in errors in the received waveforms/channel symbols. For the sake of simplicity we shall here be content to study *memoryless* channels, which transmit discrete- or continuous-valued symbols in *discrete* time intervals³⁾. Memorylessness of a channel means that the added noise at a given time instance does not influence the channel output in any *other* time instances.

6.1 Channel Coding

Channel coding means applying a certain mapping (the *channel encoder C*) from the source alphabet S to a *channel codeword alphabet C*, and a mapping from C back to S (the *channel decoder D*). In a communication system the mappings are applied as depicted in Figure 10.

The two mappings should be designed to ensure that transmission errors in the symbol sequences produced by the channel encoder will not necessarily result in errors in the decoded source symbol sequences (information messages) produced by the channel decoder. Examples of classical channel codes are algebraic forward-error-correcting codes such as *BCH codes*, and *convolutional codes* [8]. More recently the field has been revolutionized by the advent of *iterative decoding*, especially as applied to *turbo codes* and *low-density parity check* codes. We refer to the paper [9] by Øyvind Ytrehus in this issue of *Teletronikk*, and the references therein, for a tutorial introduction to these fields.

An important question now is:

Under what circumstances, if any, is it theoretically possible to design a channel coding system (encoder + decoder) such that the overall transmission of information from source to user becomes as reliable as we desire?

The answer lies in the capacity of the channel, and in *Shannon's second coding theorem*, also known as the *channel coding theorem*. In order to introduce this result, we need some notation, and a statistical model for the channels under study.

6.2 Channel Modelling

A *discrete* memoryless channel may be described by its *transition probability distribution*, i.e. a description of how probable it is that the various input symbols to the channel emerge from the channel as each of the possible output symbols. Consider a *discrete* channel having an input alphabet $\mathcal{A} = \{a_0, \dots, a_{J-1}\}$ and an output alphabet $\mathcal{B} = \{b_0, \dots, b_{K-1}\}$. A practical example is a channel where BPSK modulation symbols are transmitted ($J=2$), while the channel adds continuously distributed noise, and the receiver performs *quantization* to, say, $K=8$ levels, of the noisy (and thus amplitude-continuous) received symbols. Such quantization must be done in order to facilitate the use of digital processing during subsequent decoding.

Such a channel may be described by a *probability transition matrix P* whose (k,j) -element is P_{kj} , the probability that the channel output is b_k when the input was a_j . It is also common to visualize such a channel by an *input-output transition diagram* as exemplified in Figure 11.

For *additive-noise channels* transmitting *continuous-amplitude* symbols we may write the channel output as $Y = X + Z$ where X is the channel input and Z is the additive noise. The probability density function of the noise, $f_Z(z)$, then describes the channel.

6.3 The Channel Capacity of a Memoryless Channel

For a given memoryless channel, we now introduce the notion of *channel capacity at cost S*. The most obvious "cost" in a communication system is perhaps an upper limit on the *average symbol power* available for transmission over the channel. Another possible cost is that of bandwidth. For simplicity, we consider the case where the bandwidth is given, and thus is not a subject for optimization. The capacity at cost

³⁾ This is a valid description also for a channel transmitting continuous-time waveforms, if it is perfectly bandlimited and the Nyquist criterion is fulfilled.

S is still defined in terms of a maximum of the mutual information $I(P;Q)$ between channel input and output⁴⁾, as

$$C(S) = \max_{P \in \mathcal{P}_S} I(P;Q) \text{ bits per channel use (12)}$$

where \mathcal{P}_S is the set of all possible channel symbol distributions (discrete or continuous depending on the channel) such that the *average cost per channel use*, $E[s]$, is less than or equal to the constant S . P is our notation for an arbitrary channel symbol distribution, while Q denotes the channel's statistical model – here assumed given. Note that the probability distribution P which actually achieves the maximum is the distribution of symbols that the channel encoder must approximate if we are to achieve performance close to the capacity limit in practice.

We shall now state the channel coding theorem, which shows the physical significance of the capacity $C(S)$:

The channel coding theorem:
 Let $C(S)$ be the capacity of a memoryless channel at cost S . For any $R < C(S)$, and for any desired reliability of transmission over the channel (as measured through probability of channel decoding error) there exists a channel code of rate R (information bits per channel symbol) such that the desired reliability may be obtained. For rates $R > C(S)$ no such codes exist, and there will always be a nonzero probability of decoding errors.

What was completely novel about this result, compared to the usual line of thinking before Shannon's papers, was that it shows that *the achievable transmission rate is not a function of the desired degree of communication reliability*. Either communication *can* be made reliable, or it cannot. In the case that it can, it can be made as reliable as we want: Methods for achieving *any* desired reliability have been proven to exist for transmission rates all the way up to the (constant) channel capacity – at the cost of increased system complexity. Conversely, it has been proven that there are *no* transmission schemes to be found which can guarantee any degree of reliability if one attempts to transmit above capacity.

The channel coding theorem holds for both discrete and continuous channels. It is, perhaps, chiefly an *existence* theorem; in other words, it does not exactly tell us how to construct our channel coding systems – beyond imposing necessary constraints on the channel symbol distribution.

However, it has indeed been demonstrated that when these constraints are sought met, by so-called *shaping* [10], the system performance does improve considerably over that of the more common systems where simple uniform symbol distributions are used regardless of the channel. Also, it is worth mentioning that the high-performance turbo codes are – perhaps almost by accident, but nonetheless – constructed in a way which owes a great deal to techniques used by Shannon when proving the channel coding theorem. It may seem that information theory can be quite constructive after all, if only designers achieve enough insight into its results.

6.3.1 The Delay Problem

A limiting factor of the channel coding theorem, and indeed of many information theoretic results, is that if rates close to the capacity are desired, we may have to resort to coding extremely long blocks of symbols at a time in order to obtain the desired reliability. This will yield a very complex system with *long coding delays*. Much work has been done on so-called *delay-constrained* communication systems, and the corresponding capacity limits. The paper by Pål Orten and Bjarne Risløw [11] in this issue of *Teletronikk* contains a discussion of how additional delay constraints limit the achievable capacity.

6.4 The Capacity of a Binary Symmetric Channel

The simplest example of a communication channel one can think of is the binary symmetric channel. This channel, which models well e.g. the storage on a CD, or an optical fibre with binary modulation, takes binary symbols (denoted 0 and 1, for simplicity) as input and transmits them with a symmetric probability of error, denoted p . That is to say, the probability of a transmitted 0 being received as a 1 is p , as is the probability of a 1 being received as a 0. Note that p can be limited to $[0, 0.5]$ since, if p were higher than 0.5, the bit error rate could be improved by inverting all the received bits. The probability transition diagram for this channel is depicted in Figure 12.

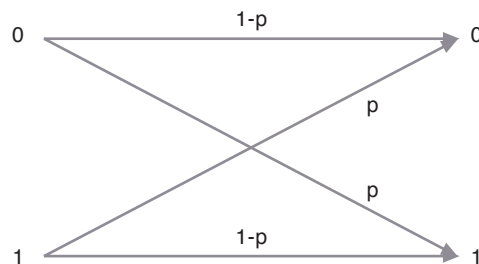


Figure 12 A binary symmetric channel

⁴⁾ The mutual information is really a function of the input distribution and the channel transition probability properties, not the random input and output variables themselves – therefore we often write it as $I(P;Q)$ instead of $I(X; Y)$.

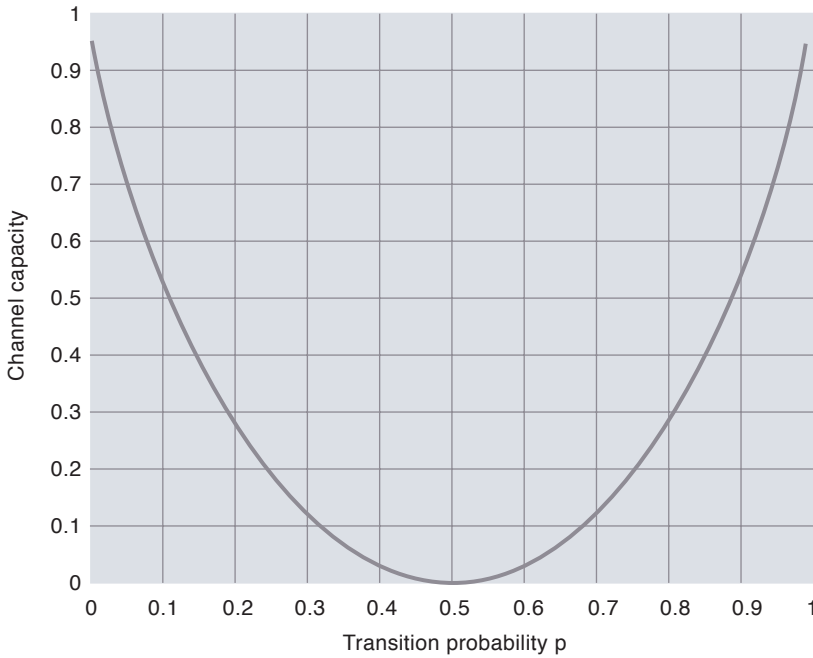


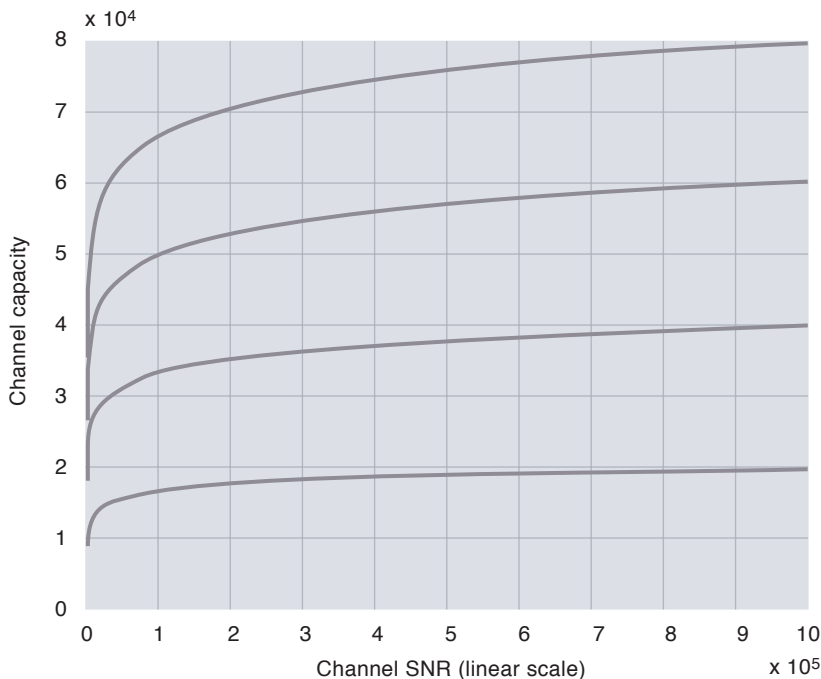
Figure 13 Channel capacity of a binary symmetric channel

The channel capacity is achieved when the input symbols are uniformly distributed, i.e. $P = \{0.5, 0.5\}$. The maximum mutual information achieved is

Figure 14 Capacity (bits per second) as a function of channel signal-to-noise ratio (SNR) for a memoryless, bandwidth- and power-limited channel with additive Gaussian noise. The capacity is depicted for bandwidths ranging from 1000 Hz (lower curve) to 4000 Hz (upper curve), in steps of 1000 Hz

$$C = 1 + p \log_2(p) + (1 - p) \log_2(1 - p) \quad \text{[information bits per channel bit]} \quad (13)$$

In other words, the capacity is $C = 1 - H(p)$, where $H(p) = -p \log_2(p) - (1 - p) \log_2(1 - p)$ is the entropy of the “channel noise source”. The higher p is (up to 0.5), the more noise there is on the channel, and the smaller the capacity is. At $p = 0.5$ there is equal probability that every received bit is correct or wrong, so there is no



chance of correct decoding. Hence, the capacity is zero. At $p = 0$ and $p = 1$ the capacity attains its maximum of 1, meaning that the channel is noiseless (in the case of $p = 1$ every single transmitted bit is inverted by the channel, but of course these “errors” can all easily be corrected by the receiver!). Thus, in these special cases no error protection (error control coding) is needed, and every transmitted channel bit can be an information bit.

6.5 The Capacity of a Gaussian Memoryless Channel

As another very important example, let us consider a memoryless continuous-amplitude channel with *additive white Gaussian zero-mean noise* (AWGN) of power (variance) N , statistically independent of the channel input. That is to say, the noise power is uniformly distributed in frequency (hence *white* noise), while the noise samples follow a Gaussian distribution.

This noise model is valid for all cases where there are many independent sources of noise; in particular it models thermal noise in electronic components well. In a cellular mobile radio system with many active users transmitting in the same frequency band, it is also a good approximation when modelling interference between users.

We furthermore assume that the channel is *bandlimited* to B Hz and that the signals to be sent are critically sampled, i.e. at $f_s = 2B$ Hz. This is a good model e.g. for a telephone line channel. For such a channel, Shannon derived the following famous formula for the capacity at received signal power S (which is the same as transmitted power if, as is usually done for this channel model, we assume that no signal attenuation is present in the channel):

$$C(S) = B \cdot \log_2 \left(1 + \frac{S}{N} \right) = B \cdot \log_2 (1 + \text{SNR}), \quad (14)$$

measured in information bits per second. SNR denotes channel *signal-to-noise ratio*; i.e. the ratio between received signal power and noise power.

This formula, whose fascinating history is traced in the paper by Lars Lundheim [12] in the present issue of *Teletronikk*, has become so well known that it can sometimes be found misinterpreted as expressing the capacity of *any* noisy channel – which it does not. It is visualized in Figure 14, for various values of the channel SNR. From the derivation, not to be done here, it is evident that if we are to actually achieve this capacity in practice, we must use a signalling alphabet with channel codewords that also fol-

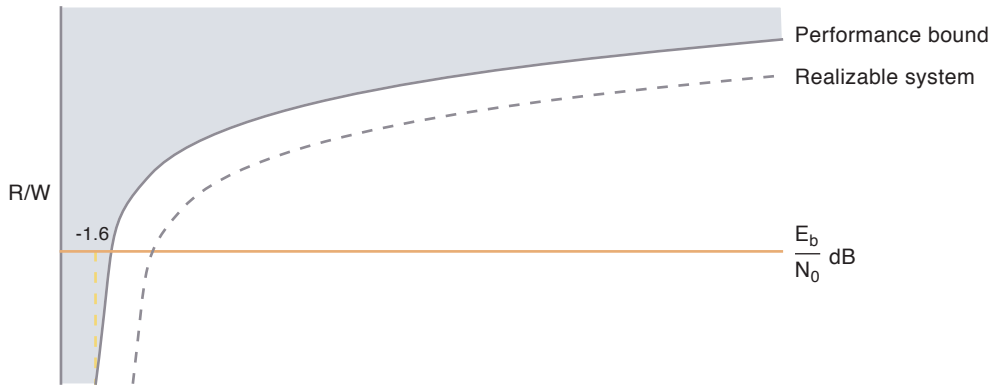


Figure 15 Capacity limit of an AWGN channel as a function of $\frac{E_b}{N_0}$

low a Gaussian distribution, with zero mean and variance S .

The capacity curves shown in Figure 14 seem to imply a linear, unlimited capacity growth as the bandwidth increases. A weakness of this figure, however, is that it does not take into account the practical fact that *thermal (white) noise power in a communication system is proportional to the system bandwidth*. Thus, operating at the same SNR at different bandwidths means that the *transmit power* is different for each bandwidth. A more “fair” comparison from a resource point of view is obtained by normalizing the SNR with respect to the bandwidth, i.e. express the SNR as the ratio of transmitted signal power S to noise power *per Hertz bandwidth*, denoted N_0 [W/Hz]. Doing this normalization, we obtain the channel capacity on an AWGN channel as follows:

$$C = B \log_2 \left(1 + \frac{S}{N_0 B} \right) \text{ [information bits/s]} \quad (15)$$

which implies a finite (assuming finite transmit power) asymptotic upper bound on the capacity as the bandwidth is increased to infinity:

$$C(B = \infty) = \frac{S}{N_0 \ln(2)}. \quad (16)$$

For a desired actual transmission rate $R \leq C$ [information bits/second] this can be used to obtain a lower bound on the *transmit energy per information bit* which must be used if error-free transmission is to be at this rate. The energy spent per channel symbol is

$$E_b = \frac{S}{R} \text{ [J/information bit]}. \quad (17)$$

Thus, $S = E_b R$, so, from Equation (15)

$$R/B \leq C/B = \log_2 \left(1 + \frac{E_b R}{N_0 B} \right), \quad (18)$$

or

$$E_b \geq N_0 \frac{2^{\frac{R}{B}} - 1}{R/B}. \quad (19)$$

The absolute lower bound for error-free transmission over the channel using *any* transmission scheme is obtained from this equation by letting the actual transmission rate go to zero:

$$\min E_b = N_0 \cdot \ln(2). \quad (20)$$

This is often expressed as the *Shannon limit* for reliable transmission on an AWGN channel,

$$\min \left\{ \frac{E_b}{N_0} \right\} = \ln(2), \quad (21)$$

or -1.6 dB on the decibel scale. Figure 15 illustrates the upper bound (18) on the achievable rate per Hz bandwidth, as a function of the signal-to-noise ratio per information bit. It can be seen that the upper rate bound goes asymptotically to zero as the SNR approaches the Shannon limit.

7 Source Compression and Rate Distortion Theory

At the beginning of this paper we studied *data compaction*, the process of efficient, *error-free* source representation. However, in many applications we are willing, or even forced, to accept some distortion in the source output – such as blurring or blocking effects in an image, or slightly “synthetic” quality, loss of treble or background noise in a speech signal – in order to fit the source data into a given channel, be that a magnetic disc, a telephone line, or a radio link.

As we have seen, each such channel is characterized by a finite *capacity* which provides an upper bound on the amount of information per channel symbol (or per unit time) which can be reliably transmitted. If we have a source whose information content is higher than the capacity of the channel we want to use, we essentially have two choices:

1. Transmit at a rate higher than the source entropy rate (and thus higher than the channel capacity). Admittedly, this enables us to send

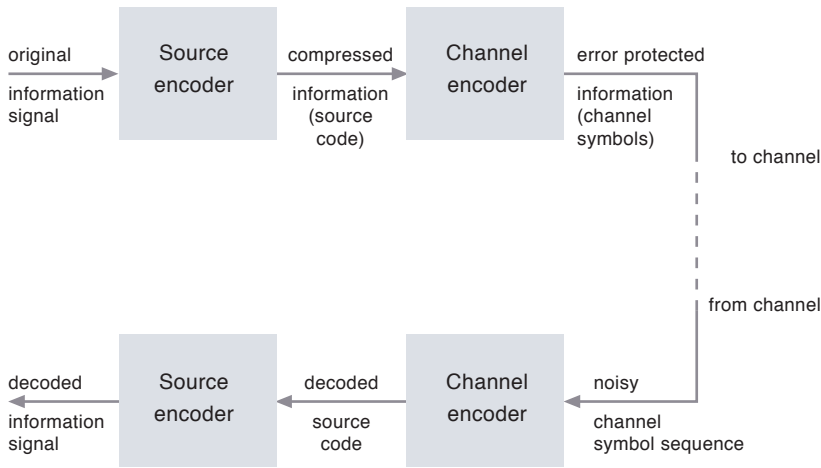


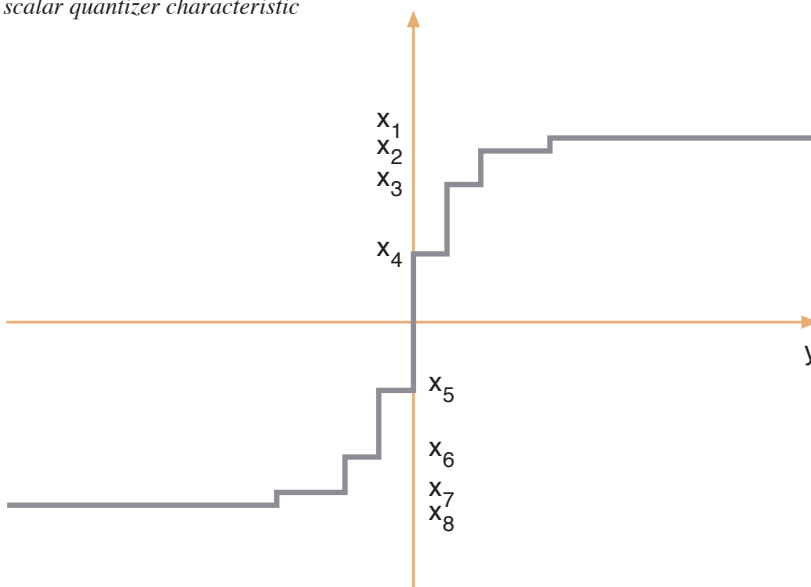
Figure 16 Information encoding and decoding operations separated into source and channel encoding and decoding

the source data error-free into the channel, but we are then forced to accept the *decoding errors* which invariably will occur after transmission over the channel. These errors will typically be outside of our control, and are stochastic in nature.

2. Before transmission, reduce the information content of the source in a controlled way (thus introducing a type of distortion over which we have some amount of control, e.g. low pass filtering or coarser amplitude quantization) until it is lower than the channel capacity. The quality-reduced but hopefully still usable source can then in principle be reliably transmitted over the given channel.

Developing good methods for implementing choice 2 is the task of *source compression*. In general, source compression is performed by a *source coder-decoder pair*, which is placed

Figure 17 3-bit non-uniform scalar quantizer characteristic



within a complete communication system as shown in Figure 16.

The idea that the tasks of source and channel coding can be separated as shown above, without losing overall system optimality, but greatly reducing design complexity, is rooted in the *separation principle*, originally devised by Shannon. However, it is important to realize that this principle holds only when there are *no constraints on computational complexity or overall delay* in the system. For real systems where such practical constraints are imposed, there might be performance advantages in a joint design of source and channel coding. The paper by Tor A. Ramstad [13] in this *Teletronikk* issue addresses this problem.

7.1 Scalar Quantization

The very simplest way of compressing a source's information content is that of *scalar amplitude quantization*. A scalar quantizer Q is a nonlinear operation which is used, on a single-symbol basis, to limit the source alphabet to a finite, countable set – i.e. for each real-valued input symbol x , which may be taken from a continuous distribution, the quantizer outputs one of only a limited number N of amplitude levels or *reproduction values*. The output $y = Q(x)$ is in each case chosen based on a *nearest neighbour rule*. That is, for each input symbol, the quantizer simply outputs the quantized output level which is closest to the input on the real line. It is common to visualize a scalar quantizer by means of its *quantizer characteristic*, which simply depicts the input-output relation $y = Q(x)$ as a function.

In Figure 17 a *non-uniform* scalar quantizer characteristic is shown. A quantizer is said to be non-uniform when the quantization *intervals* – the intervals between the possible output intervals – have varying length. This is beneficial when quantizing sources with non-uniform pdfs. One important example is in the CCITT A-law used for speech in Pulse Code Modulation (PCM) for telephony [14]. Shown here is a *3-bit* quantizer, which simply means that there are $8 = 2^3$ output levels, which can be uniquely indexed by means of 3-bit strings or binary codewords. Thus the information content of the quantized source in this case is not more than 3 bits per source symbol, whereas the information content before quantization was infinitely large.

7.2 Rate Distortion Theory

The branch of information theory that provides the foundation for source compression – i.e. the bounds we search to reach when we compress e.g. an image or an audio signal – is called *rate distortion theory*. This theory is first and foremost concerned with the following question:

For a given source S , and a given representation rate R [code symbols/source symbol], find the minimal distortion D that can theoretically be attained in the reproduction of S – or, equivalently: For a given maximal distortion D we are willing to accept in the reproduction of S , what is the minimal rate R which is theoretically possible to use?

The answer to the above question(s) is given by the *rate distortion* (or *source distortion*) function $R(D)$, which for a given (discrete or continuous) source S is also defined in terms of the mutual information function:

$$R(D) = \min_{Q \in Q_D} I(S; Q). \quad (22)$$

Here Q_D is the set of all possible noisy *test channels* that yields an average symbol distortion $E[d]$ less than or equal to D when the source S is transmitted through the channel.

That is, one imagines that the distortion due to compression of the source has occurred due to the transmission of the source over some (imaginary) noisy channel, described by a transition probability matrix or noise distribution – a channel that gives rise to a distortion not larger than D . All channels for which this is possible to achieve are considered; then one chooses the one channel where the lowest transmission rate could be used. This minimal rate is then $R(D)$, and the noise characteristics of the corresponding channel tell us how the distortion due to coding should be distributed among the source symbols.

The designer challenge when designing good compression algorithms is now to find an actual algorithm which results in reproduction noise with the same noise characteristics as those of the imaginary channel mentioned above.

The choice of *symbol distortion measure*, d , is arbitrary and may be done by the designer. In practice, one often uses the *mean square* distortion between original and decompressed source symbols, i.e. $d = (X - Y)^2$, such that $E[(X - Y)^2] \leq D$. This is mainly due to the computational tractability and simplicity of this measure.

The exact physical significance of the rate distortion function is summed up in *Shannon's third coding theorem*, also termed the *source-compression theorem*:

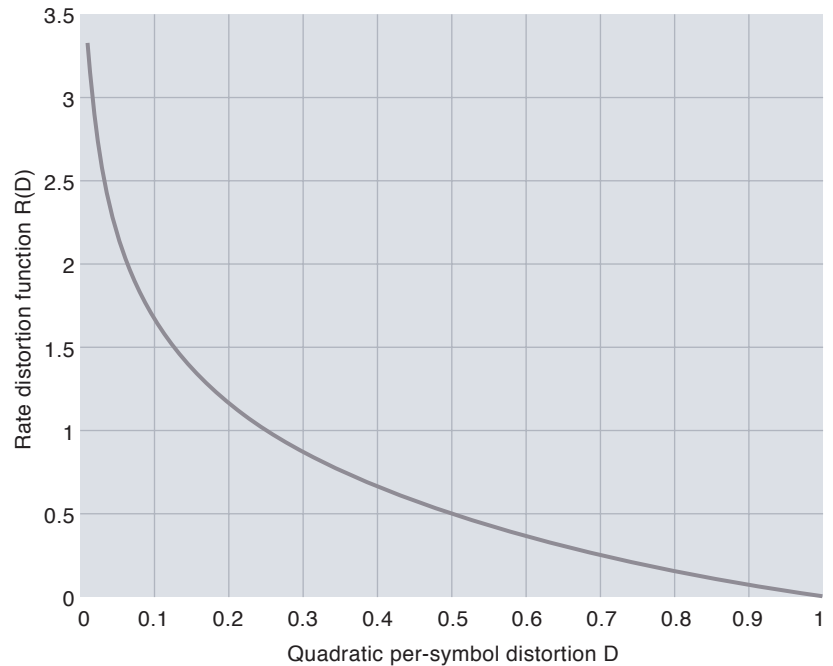


Figure 18 The rate-distortion function for a memoryless $N(0, 1)$ source

The source-compression theorem

Let S be a memoryless source with rate distortion function $R(D)$. Then, for every $D > 0$ it is possible to find a source code of rate $r > R(D)$ such that the source after being encoded and decoded with this code is reproduced with average symbol distortion less than or equal to D . If $r \leq R(D)$ there exists no code such that this is possible.

Again, the theorem holds for both discrete and continuous sources. It does not tell us exactly *how to find* practically realizable source codes that approach the rate-distortion function for a given source. However, from the proof of the theorem, as for the case of channel coding, it is evident that we may have to consider encoding symbol blocks of possibly infinite length n (and hence infinite delay and coder complexity) in order to do this.

As an example, consider a continuous, memoryless source with symbol outcomes following a *Gaussian* distribution $N(0, \sigma^2)$. This is a reasonably valid model e.g. for sub-band samples coming from a well-designed *analysis filterbank* [15]. For such a source, the rate distortion function is given by

$$R(D) = \frac{1}{2} \log_2 \frac{\sigma^2}{D}, \text{ for } 0 \leq D \leq \sigma^2 \quad (23)$$

In Figure 18 this function is depicted, for $\sigma^2 = 1$. Even though closed form expressions for the rate distortion function may be found for only relatively simple source models [6], its basic general properties may be generally derived from the formal definition:

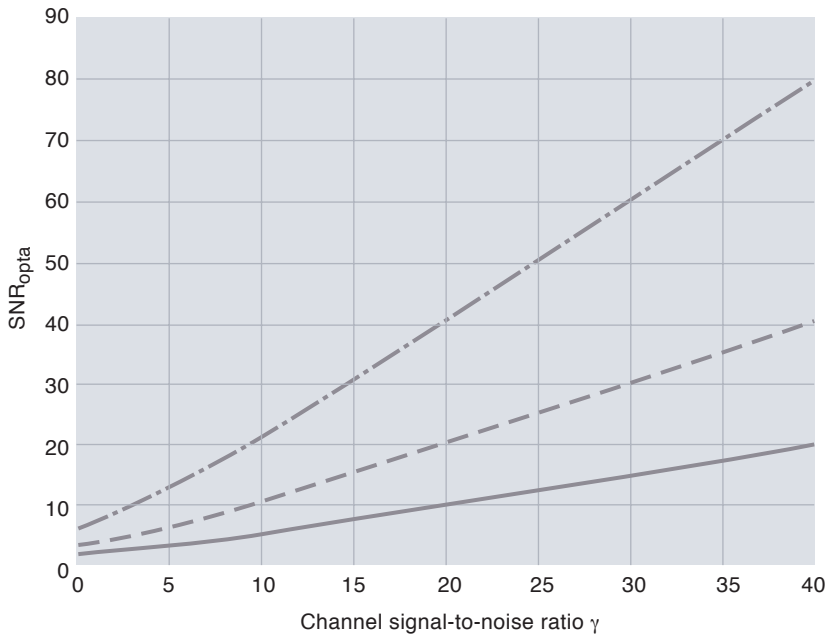


Figure 19 Optimal signal-to-noise ratio (dB) after decoding, as a function of the channel signal-to-noise ratio (dB), for a Gaussian memoryless source transmitted over a memoryless, bandlimited AWGN channel. Solid line: $T_s = T_c/2$ (bandwidth reduction during transmission). Dashed line: $T_s = T_c$. Dash-dotted line: $T_s = 2T_c$ (bandwidth expansion during transmission)

The function always goes from the rate value $H(S)$ (which here is infinite, since the source is continuous) at distortion 0, to rate 0 at some finite maximal distortion D_{\max} . Between these two points the function is always strictly decreasing and convex; hence also continuous. This implies that we can be sure that if we increase the rate, it is always possible to obtain a source representation that has strictly improved, at least in a signal-to-noise ratio sense. The relative improvement will be largest at low rates.

7.3 OPTA – Optimum Performance Theoretically Attainable

The capacity bound for noisy channels and the rate distortion bound for compression may be combined to produce what is commonly referred to as the *Optimum Performance Theoretically Attainable*, or *OPTA*. This is the best performance that can ever be achieved in a system where a certain source is to be transmitted over a given channel. If the rate distortion function of the source is $R(D)$, and the channel capacity is C , the minimal distortion that can be achieved is found by solving the equation $C = R(D)$ with respect to D .

For an AWGN channel where the channel symbol rate is $1/T_c$ [channel symbols/s] and the source symbol rate is $1/T_s$ [source symbols/s], this solution can be found on a simple closed form [13]:

$$D_o = S \left(1 + \frac{S}{N_0 B} \right)^{-\frac{T_s}{T_c}}, \quad (24)$$

with corresponding maximal SNR after decoding given by

$$\text{SNR}_o = \frac{S}{D_o}. \quad (25)$$

In Figure 19 this optimally achievable SNR after decoding is shown, for various ratios between the source and channel symbol rate. It is seen that the more channel bandwidth spent (bandwidth expansion), the better the fidelity. This is because bandwidth expansion allows for the use of more powerful error control schemes.

8 Concluding Remarks

This paper has merely scratched the surface of a huge and continuously expanding field, and tried to convey some basic insights into how information theory works. There is a lot more to be gained from studies of information theoretic topics than what could possibly be included here, e.g. error analysis of communication systems, construction and performance analysis of channel and source codes, optimization of quantization schemes, and so on.

Information theory does of course not provide us with all the answers to questions regarding communication system design. However, it does often provide us with important pointers as to what we should or should not do. For an example, consider modern telephone line modem standards utilizing *trellis coded modulation* (TCM) techniques [16], with throughput rates up to 33.4 kbit/s. 15 years ago, it was thought “impossible” to implement such modem rates in practice due to channel noise problems. Yet the capacity theorems of information theory have all the time indicated that the fundamental attainable rate limit for error-free transmission is far higher — approximately 60 kbit/s for today’s telephone lines. The modern high-rate modems would have been unimaginable without the insights that information theory provides, as would the knowledge of the actual potential to be gained.

The same can be said for most modern channel coding, modulation, compression, and compaction techniques. It is thus clear that information theory can be of much more practical use than what is suggested by the maybe too-common picture many communication engineers have of the field, as a set of theoretical bounds, attainable only with the aid of both unknown, and possibly infinitely complex, algorithms. It is a challenge to those of us who are teaching information theoretic concepts to replace this picture with a more positive and useful one.

Finally, we remark that there are still many challenges for information theory, and communication system design problems which might benefit from its insight. We would particularly like to mention the following fields as exciting areas for information theoretic research in the future:

- Design and performance analysis of multiuser communication networks and MAC protocols;
- Time-varying and frequency-dispersive wireless channels, particularly multiple-input-multiple-output (MIMO) channels;
- Diversity issues in general.

It seems quite clear that information theory still has an important role to play if tomorrow's communication systems are to live up to the expectations towards them. Shannon's ideas are likely to cast long shadows into the 21st century.

References

- 1 Shannon, C E. A mathematical theory of communication. *Bell Syst. Tech. J.*, 27, 379–423 and 623–656, 1948.
- 2 Shannon, C E. Communication in the presence of noise. *Proc. IRE*, 37, 10–21, Jan. 1949.
- 3 Shannon, C E. Coding theorems for a discrete source with a fidelity criterion. *IRE Nat. Conv. Rec.*, 142–163, Mar. 1959.
- 4 Blahut, R E. *Principles and Practice of Information Theory*. Reading, MA, Addison Wesley, 1987.
- 5 Cover, T M, Thomas, J A. *Elements of Information Theory*. New York, Wiley, 1991.
- 6 Berger, T. *Rate Distortion Theory*. Englewood Cliffs, NJ, Prentice-Hall, 1971.
- 7 Huffman, D A. A method for the construction of minimum redundancy codes. *Proc. IRE*, 40, 1098–1101, 1952.
- 8 Blahut, R E. *Theory and Practice of Error Control Codes*. Addison-Wesley, 1984.
- 9 Ytrehus, Ø. An Introduction to Turbo Codes and Iterative Decoding. *Teletronikk*, 98 (1), 65–77, 2002. (This issue.)
- 10 Forney, G D Jr. Trellis shaping. *IEEE Trans. on Information Theory*, 38 (2), 1992.
- 11 Orten, P, Risløw, B. Theory and Practice of Error Control Coding for Satellite and Fixed Radio Systems. *Teletronikk*, 98 (1), 78–91, 2002. (This issue.)
- 12 Lundheim, L. On Shannon and “Shannon's formula”. *Teletronikk*, 98 (1), 20–29, 2002. (This issue.)
- 13 Ramstad, T. Shannon Mappings for Robust Communication. *Teletronikk*, 98 (1), 114–129, 2002. (This issue.)
- 14 Haykin, S. *Communication Systems*. Wiley, 1994 (3rd ed.).
- 15 Ramstad, T A, Aase, S O, Husøy, J H. *Subband Compression of Images – Principles and Examples*. North Holland, Elsevier, 1995.
- 16 Biglieri, E et al. *Introduction to Trellis-coded modulation with applications*. New York, Macmillan, 1991.

On Shannon and “Shannon’s Formula”

LARS LUNDHEIM



Lars Lundheim (44) received his *Siv.ing.* and *Dr.ing.* degrees from the Norwegian University of Science and Technology (NTNU), Trondheim, in 1984 and 1992, respectively. He has held various research and teaching positions at NTNU, SINTEF, CERN and Trondheim College of Engineering. He is currently employed as Associate Professor at NTNU, doing research and teaching in signal processing for wireless communication.

lundheim@tele.ntnu.no

The period between the middle of the nineteenth and the middle of the twentieth century represents a remarkable period in the history of science and technology. During this epoch, several discoveries and inventions removed many practical limitations of what individuals and societies could achieve. Especially in the field of communications, revolutionary developments took place such as high speed railroads, steam ships, aviation and telecommunications.

It is interesting to note that as practical limitations were removed, several fundamental or principal limitations were established. For instance, Carnot showed that there was a fundamental limit to how much energy could be extracted from a heat engine. Later this result was generalized to the second law of thermodynamics. As a result of Einstein’s special relativity theory, the existence of an upper velocity limit was found. Other examples include Kelvin’s absolute zero, Heisenberg’s uncertainty principle and Gödel’s incompleteness theorem in mathematics. Shannon’s Channel coding theorem, which was published in 1948, seems to be the last one of such fundamental limits, and one may wonder why all of them were discovered during this limited time-span. One reason may have to do with maturity. When a field is young, researchers are eager to find out what can be done – not to identify borders they cannot pass. Since telecommunications is one of the youngest of the applied sciences, it is natural that the more fundamental laws were established at a late stage.

In the present paper we will try to shed some light on developments that led up to Shannon’s information theory. When one compares the generality and power of explanation of Shannon’s paper “A Mathematical Theory of Communication” [1] to alternative theories at the time, one can hardly disagree with J.R. Pierce who states that it “came as a bomb” [4]. In order to see the connection with earlier work, we will therefore focus on one particular case of Shannon’s theory, namely the one which is sometimes referred to as “Shannon’s formula”. As will be shown, this result was discovered independently by several researchers, and serves as an illustration of a scientific concept whose time had come. Moreover, we will try to see how development in this field was spurred by techno-

logical advances, rather than theoretical studies isolated from practical life.

Besides the original sources cited in the text, this paper builds on historical overviews, such as [4] and [19]–[23].

“Shannon’s Formula”

Sometimes a scientific result comes quite unexpected as a “stroke of genius” from an individual scientist. More often a result is gradually revealed, by several independent research groups, and at a time which is just ripe for the particular discovery. In this paper we will look at one particular concept, the channel capacity of a band-limited information transmission channel with additive white, Gaussian noise. This capacity is given by an expression often known as “Shannon’s formula¹⁾”:

$$C = W \log_2(1 + P/N) \text{ bits/second.} \quad (1)$$

We intend to show that, on the one hand, this is an example of a result for which time was ripe exactly a few years after the end of World War II. On the other hand, the formula represents a special case of Shannon’s information theory²⁾ presented in [1], which was clearly ahead of time with respect to the insight generally established.

“Shannon’s formula” (1) gives an expression for how many bits of information can be transmitted without error per second over a channel with a bandwidth of W Hz, when the average signal power is limited to P watt, and the signal is exposed to an additive, white (uncorrelated) noise of power N with Gaussian probability distribution. For a communications engineer of today, all the involved concepts are familiar – if not the result itself. This was not the case in 1948. Whereas bandwidth and signal power were well-established, the word *bit* was seen in

1) Many mathematical expressions are connected with Shannon’s name. The one quoted here is not the most important one, but perhaps the best known among communications engineers. It is also the one with the most immediately understandable significance at the time it was published.

2) For an introduction to Shannon’s work, see the paper by N. Knudtzon in this issue.

print for the first time in Shannon's paper. The notion of probability distributions and stochastic processes, underlying the assumed noise model, had been used for some years in research communities, but was not part of an ordinary electrical engineer's training.

The essential elements of "Shannon's formula" are:

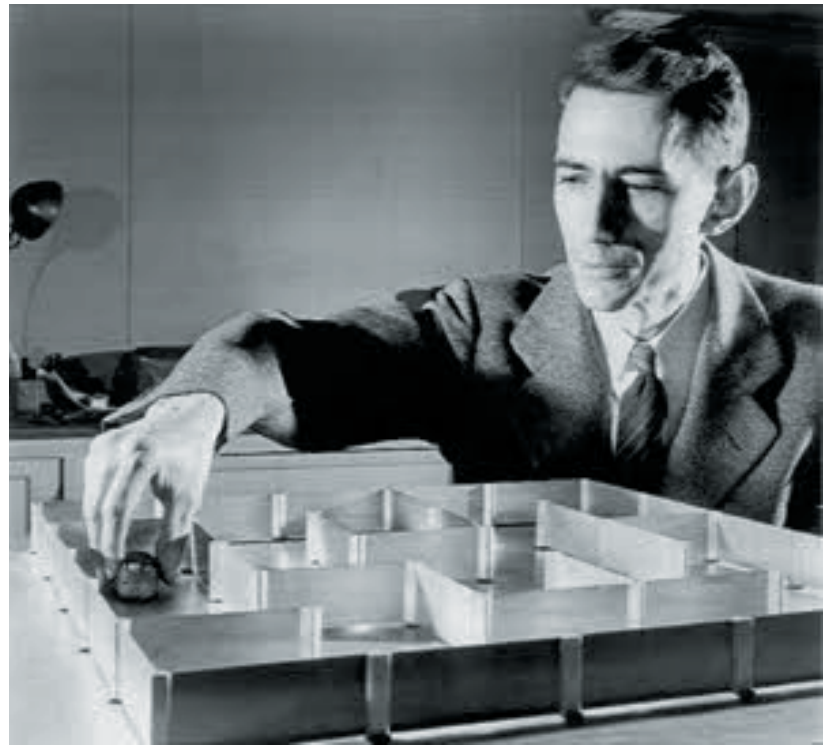
1. Proportionality to bandwidth W
2. Signal power S
3. Noise power P
4. A logarithmic function

The channel bandwidth sets a limit to how fast symbols can be transmitted over the channel. The signal to noise ratio (P/N) determines how much information each symbol can represent. The signal and noise power levels are, of course, expected to be measured at the receiver end of the channel. Thus, the power level is a function both of transmitted power and the attenuation of the signal over the transmission medium (channel).

The most outstanding property of Shannon's papers from 1948 and 1949 is perhaps the unique combination of generality of results and clarity of exposition. The concept of an information source is generalized as a symbol-generating mechanism obeying a certain probability distribution. Similarly, the channel is expressed essentially as a mapping from one set of symbols to another, again with an associated probability distribution. Together, these two abstractions make the theory applicable to all kinds of communication systems, man-made or natural, electrical or mechanical.

Independent Discoveries

One indicator that the time was ripe for a fundamental theory of information transfer in the first post-war years is given in the numerous papers attempting at such theories published at that time. In particular, three sources give formulas quite similar to (1). The best known of these is the book entitled *Cybernetics* [2] published by Wiener in 1949. Norbert Wiener was a philosophically inclined and proverbially absent-minded professor of mathematics at MIT. Nonetheless, he was deeply concerned about the application of mathematics in all fields of society. This interest led him to founding the science of cybernetics. This field, which is perhaps best defined by the subtitle of [2]: "Control and Communication in the Animal and the Machine" included, among other things, a theory for infor-



Claude Elwood Shannon (1916–2001), the founder of information theory, also had a practical and a playful side. The photo shows him with one of his inventions: a mechanical "mouse" that could find its way through a maze. He is also known for his electronic computer working with Roman numerals and a gasoline-powered pogo stick

mation content in a signal and the transmission of this information through a channel. However, Wiener was not a master of communicating his ideas to the technical community, and even though the relation to Shannon's formula is pointed out in [2], the notation is cumbersome, and the relevance to practical communication systems is far from obvious.

Reference to Wiener's work was done explicitly by Shannon in [1]. He also acknowledged the work by Tuller³⁾. William G. Tuller was an employee at MIT's Research Laboratory for Electronics in the second half of the 1940s. In 1948 he defended a thesis at MIT on "Theoretical Limitations on the Rate of Transmission of Information"⁴⁾. In his thesis Tuller starts by referring to Nyquist's and Hartley's works (see below). Leaning on the use of sampling and quantization of a band-limited signal, and arguing that intersymbol interference introduced by a band-limited channel can in principle be eliminated, he states quite correctly that under noise-free conditions an unlimited amount of information can be transmitted over such a channel. Taking noise into account, he delivers an argument partly based on intuitive reasoning, partly on formal mathematics, arriving at his main result that the information H transmitted over a transmission link of bandwidth B during a time interval T with carrier-to-noise-ratio C/N is limited by

³⁾ Shannon's work was in no way based on Wiener or Tuller; their then unpublished contributions had been pointed out to Shannon after the completion of [1].

⁴⁾ Later published as RLE Technical report 114 and as a journal paper [5] (both in 1949).

Norbert Wiener (1894–1964) had been Shannon’s teacher at MIT in the early 1930s. By his seminal work *Extrapolation, Interpolation and Smoothing of Stationary Time Series* made during World War II he lay the foundation for modern statistical signal processing. Although Shannon was influenced by Wiener’s ideas, they had little or no contact during the years when they made their contributions to communication theory. Their styles were very different. Shannon was down-to-earth in his papers, giving illustrative examples that made his concepts possible to grasp for engineers, and giving his mathematical expression a simple, crisp flavour. Wiener would rather like to use the space in-between crowded formulas for philosophical considerations and esoteric topics like Maxwell’s demon



$$H \leq 2BT \log(1 + C/N). \quad (2)$$

This expression has a striking resemblance to Shannon’s formula, and would by most readers be considered equivalent. It is interesting to note that for the derivation of (2) Tuller assumes the use of PCM encoding.

A work not referenced by Shannon is the paper by Clavier [16]⁵⁾. In a similar fashion to Tuller, starting out with Hartley’s work, and assuming the use of PCM coding, Clavier finds a formula essentially equivalent to (1) and (2). A fourth independent discovery is the one by Laplume published in 1948 [17].

Early Attempts at a General Communication Theory

Shannon and the other researchers mentioned above were not the first investigators trying to find a general communication theory. Both Shannon, Tuller and Clavier make references to the work done in the 1920s by Nyquist and Hartley.

By 1920 one can safely say that telegraphy as a practical technological discipline had reached a mature level. Basic problems related to sending and receiving apparatus, transmission lines and cables were well understood, and even wireless transmission had been routine for several years. At this stage of development, when only small increases in efficiency are gained by technological improvements, it is natural to ask whether one is close to fundamental limits, and to try to understand these limits. Harry Nyquist, in his paper “Certain Factors Affecting Telegraph Speed” [7], seems to be the first one to touch

upon, if not fully identify, some of the issues that were clarified by Shannon twenty years later.

First, it is obvious to Nyquist that the “Speed of transmission of intelligence” (which he terms W) is limited by the bandwidth of the channel⁶⁾.

Without much formal mathematical argument, Nyquist derives the following approximate formula for W :

$$W = K \log m \quad (3)$$

where m is the “number of current values”, which in modern terms would be called “the size of the signalling alphabet” and K is a constant.

Whereas Nyquist’s paper is mostly concerned with practical issues such as choice of pulse waveform and different variants of the Morse code, a paper presented three years later by Hartley is more fundamental in its approach to the problem. The title is simply “Transmission of Information”, and in the first paragraph the author says that “What I hope to accomplish (...) is to set up a quantitative measure whereby the capacities of various systems to transmit information may be compared”. Even though Nyquist had given parts of the answer in his 1924 paper, this is the first time the question that was to lead up to Shannon’s information theory is explicitly stated.

Compared to Nyquist, Hartley went a couple of steps further. For one thing, he stated explicitly that the amount of information that may be transmitted over a system⁷⁾ is proportional to the bandwidth of that system. Moreover, he formulated what would later be known as *Hartley’s law*, that information content is proportional to the product of time (T) and bandwidth (B), and that one quantity can be traded for the other. It should also be mentioned that Hartley argued that the theory for telegraph signals (or digital signals in modern terms) could be generalized to continuous-time signals such as speech or television. Hartley’s law can be expressed as

$$\text{Amount of information} = \text{const} \cdot BT \cdot \log m. \quad (4)$$

Relations between bandwidth and time similar to the one found by Nyquist was discovered simultaneously by Karl Küpfmüller in Germany [10].

⁵⁾ It is, perhaps, strange that neither Shannon nor Clavier have mutual references in their works, since both [3] and [16] were orally presented at the same meeting in New York on December 12, 1947, and printed more than a year afterwards.

⁶⁾ Proportionality to bandwidth is not explicitly stated by Nyquist in 1924. He has probably been aware of it, and includes it in his more comprehensive paper [7] four years later.

⁷⁾ Neither Nyquist nor Hartley make explicit distinction between source, channel and destination, as Shannon does twenty years later. This may seem like a trivial omission, but the distinction is essential for Shannon’s general definition of channel capacity, which requires this separation to define quantities such as source entropy and mutual information.

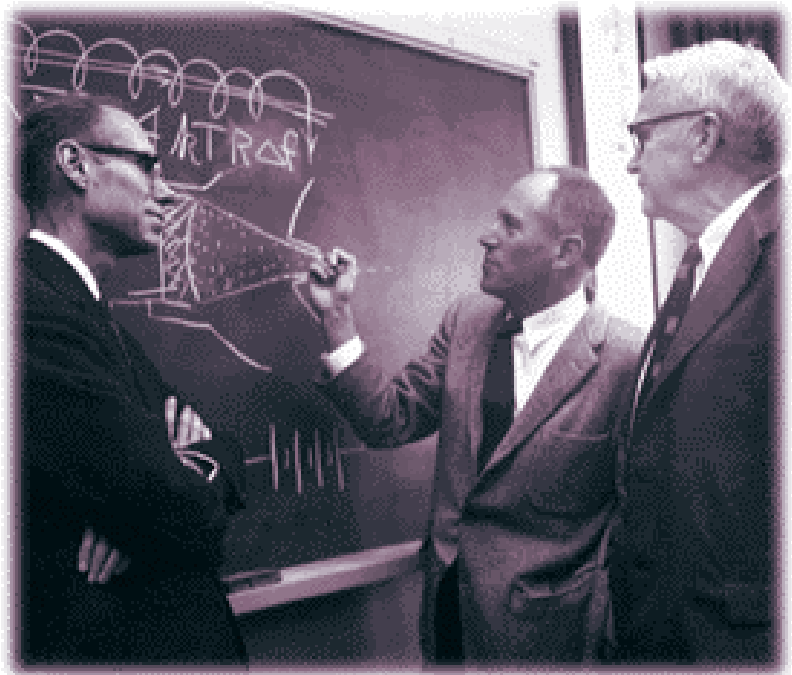
A more mathematically stringent analysis of the relation was carried out in Gabor's "Theory of communication".

As was pointed out by Tuller [5], a fundamental deficiency of the theories of Nyquist, Hartley, Küpfmüller and Gabor, is that their formulas do not include noise. The role of noise is that it sets a fundamental limit to the number of levels that can be reliably distinguished by a receiver. From expressions (3) and (4) we see that both Nyquist and Hartley were aware of the fact that the amount of information depends on the number of distinguishable signal levels (or symbols). However, they seem content to include this number m in their formulas instead of deriving it from a more fundamental quantity, such as the signal-to-noise level. In a short discussion, Nyquist mentions "interference" as one of the limiting factors of the signal alphabet size. Hartley points out the inter-symbol interference due to channel distortion as the most important limiting factor. This is fundamentally wrong, as Tuller remarks, since inter-symbol interference can, in principle, be removed by an equalizer. This is precisely what Nyquist shows in his 1928 paper [7].

Developments in the Understanding of Bandwidth

As we have seen, the work of Nyquist, Hartley and Küpfmüller in the 1920s represents an important step towards the fully developed channel capacity formulas expressed twenty years later. A key insight was the realization that information transmission rate was limited by system bandwidth. We have argued that this understanding came when the maturity of telegraph technology made it natural to ask if fundamental limitations were within reach. Another, related factor was that the concept of bandwidth, so essential in the cited works, was now clearly understood by the involved researchers. Today, when students of electrical engineering are exposed to Fourier analysis and frequency domain concepts from early on in their education, it seems strange that such a fundamental signal property as bandwidth was not fully grasped until around 1920. We will therefore take a look at how frequency domain thinking was gradually established as communications technology evolved.

If one should assign a birth date to practical (electrical) telecommunications technology, it would be natural to connect it to one of the first



successful demonstrations of telegraphy by Veil and Morse in the 1840s. It took only a few years before all developed countries had established systems for telegraph transmission. These systems were expensive, both to set up and to operate, and from the start it was important to make the transmission as cost-effective as possible. This concern is already reflected in the Morse code⁹⁾ alphabet which is designed according to the relative frequencies of characters in written English.

It was soon evident that, for transmission on overhead wires, the transmission speed was limited by the telegraph operator and, possibly, the inertia of the receiving apparatus, not by any properties of the transmission medium. The only problem connected to what we today would call the channel was signal attenuation. This was, however, a relatively small problem, since signal retransmission was easily accomplished, either manually or with automatic electromagnetic repeater relays.

For the crossing of rivers or stretches of ocean, the telegraph signals were transmitted using cables. This transmission medium soon showed to be far more problematic than overhead lines. For long spans repeaters were not a practical solution, meaning that the attenuation problem became serious. It was also found that operators

Harry Nyquist (right) (1889–1976) with John R. Pierce (left) and R. Kompfner. Nyquist was born in Nilsby in Värmland, Sweden, and emigrated to USA in 1907⁸⁾. He earned an MS degree in Electrical Engineering in 1914 and a PhD in physics at Yale in 1917. The same year he was employed by AT&T where he remained until his retirement in 1954. Nyquist made contribution in very different fields, such as the modelling of thermal noise, the stability of feed-back systems, and the theory of digital communication

⁸⁾ More about Harry Nyquist's early years can be found on Lars-Göran Nylén's homepage, URL: <http://members.tripod.com/~lgn75/>

⁹⁾ The Vail Code would be a more proper term, since it was Morse's assistant Alfred Vail who in 1837 visited a print shop in Morristown to learn from the contents of the type cases which letters were more frequent in use. He then advised Morse to abandon his plan of using a word code involving the construction of a dictionary assigning a number to all English words, and use the more practical character code with unique dash-dot combinations for each letter [19].

Fig. 5.



Detail from Alexander Graham Bell's patent for the "Harmonic Telegraph". An alternating current is generated in the left coil with a frequency given by the vibrating reed *c*. The reed *h* to the right will resonate and give a sound only if it is tuned to the same frequency. By this mechanism several users could transmit telegraph signal over the same line by using equipment tuned to different frequencies

would have to restrain themselves and use a lower speed than normal to obtain a distinguishable message at the receiver end. Both these problems were of concern in the planning of the first transatlantic telegraph cable. Expert opinions were divided from the start, but the mathematical analysis of William Thomson (later Lord Kelvin) showed that even though the attenuation would be large, practical telegraphy would be possible by use of sensitive receiving equipment. In particular, Thomson's analysis explained how the dispersion of the cable sets a limit to the possible signalling speed.

In our connection, Thomson's work is interesting because it was the first attempt of mathematical analysis of a communication channel. We see that two of the four elements of Shannon's formula were indirectly taken into account: signal power reduced by the attenuation and bandwidth limiting the signalling speed. Bandwidth was not explicitly incorporated in Thomson's theory. This is quite natural, since the relevant relationships were expressible in physical cable constants such as resistance and capacitance. These were parameters that were easily understood and that could be measured or calculated. Bandwidth, on the other hand, was simply not a relevant notion, since engineers of the time had not yet learnt to express themselves in frequency domain terms.

During the nineteenth century topics such as oscillation, wavelength and frequency were thoroughly studied in fields such as acoustics, optics and mechanics. For electrical and telegraph engineers, however, these concepts had little interest from the start.

Resonance was a well-known phenomenon in acoustics. It was therefore an important conceptual break-through when Maxwell showed mathematically how a circuit containing both capacitance and inductance would respond significantly different when connected to generators producing alternating current of different frequencies. The phenomenon of electrical reso-

nance was then demonstrated in practical experiments by Hertz¹⁰.

It is interesting to see how, even before electrical resonance was commonly understood, acoustical resonance was suggested as a means of enhancing the capacity of telegraph systems. As noted above, the telegraph operator represented the bottleneck in transmission by overhead wires. Thus, many ingenious schemes were suggested by which two or more operators could use the same line at a time. As early as 1853 the American inventor M.B. Farmer is reported to have suggested the first system for time division multiplex (TDM) telegraphy. The idea, which was independently set forth several times, was perfected and made practical by the Frenchman J.M.E. Baudot¹¹ around 1878. In parallel with the TDM experiments, several inventors were working with frequency division multiplex (FDM) schemes. These were based on vibrating reeds, kept in oscillation by electromagnets. By assigning a specific frequency to each telegraph operator, and using tuned receivers, independent connections could be established over a single telegraph line. One of the most famous patents is the "harmonic telegraph" by A.G. Bell from the 1870s. It was during experiments with this idea that Bell more or less accidentally discovered a way of making a practical telephone.

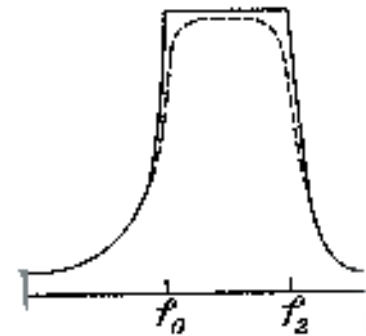
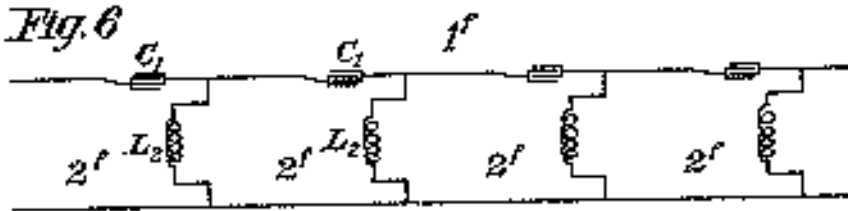
By 1890 electrical resonance phenomena were generally understood by scientists and well-informed engineers. Consequently, practical patents on how to use electrical resonance in FDM telegraphy began to be filed. These attempts, which also included some ideas of FDM telephony (not practical at the time) were, however, overshadowed by a new invention: *the wireless*.

After the first few experiments with wireless telegraphy at the start of the twentieth century, it became clear that sharp resonance circuits, tuned to specific frequencies or wavelengths were necessary to avoid disturbance among different users. This requirement made it necessary for radio engineers to have a good understanding of frequency and the behaviour of electrical circuits when connected to sources of varying frequency content. It is important to note that the concept of *bandwidth* was much more elusive than that of frequency. Today, the decomposition of an information-bearing signal into its Fourier components is a routine operation to any electrical or communications engineer. This was not the case 80–100 years ago. At that time, one would commonly assume that a radio transmitter was tuned

¹⁰) For more details of the history of electrical resonance, including an early contribution by Thomson, see the paper by Blanchard [20].

¹¹) Baudot is today remembered by the unit baud for measuring the number of symbols per second transmitted through a communication channel.

Fig. 6



Example of Campbell's bandpass filter designs with transfer function. (Original figures from U.S. Patent No 1 227 113)

to one – and only one – frequency. The bandwidth of a telegraph signal was small compared to both the carrier frequencies used and to the width of the resonance circuits employed. A general awareness of bandwidth did not develop until some experience had been gained with telephone transmission.

In parallel with the development of wireless telegraphy, and, gradually, telephony, work continued on FDM in telecommunications systems or “carrier current telephony and telegraphy”¹²⁾ which was the term used at the time. In these systems, first intended for enhancing the capacity of long-distance wire-bound connections, it became important to minimize the spacing between carrier frequencies and at the same time prevent cross-talk between the channels. To obtain this, traditional resonance circuits were no longer adequate for transmitter or for receiver tuning. What was needed was band-pass filters, sufficiently broad to accept the necessary bandwidth of the modulated speech signal, flat enough to avoid distortion, and with sufficient stop-band attenuation to avoid interference with neighbour channels. This kind of device was developed during the years prior to World War I, and was first patented by G.A. Campbell of Bell Systems in 1917.

With hindsight it is curious to note that before Campbell's invention, band-limited channels in a strict and well-defined way did not exist! Earlier transmission channels were surely band-limited in the sense that they could only be used in practice for a limited frequency range. However, the frequency response tended to roll-off gradually so as to make the definition of bandwidth, such as is found in Shannon's formula, questionable.

So, around 1920 it was evident that an information-bearing signal needed a certain bandwidth,

whether it was to be transmitted in original or modulated (by a carrier) form. There was, however, some discussion as to how large this bandwidth had to be. The discussion seems to have ceased after Carson's “Notes on the Theory of Modulation” in 1922. By this time it had been theoretically shown (by Carson) and practically demonstrated that by using so-called single side-band modulation (SSB) a modulated signal can be transmitted in a bandwidth insignificantly larger than the bandwidth of the original (unmodulated) signal. The aim of Carson's paper was to refute claims that further bandwidth reduction could be obtained by modulating the frequency of the carrier wave instead of the amplitude¹³⁾. An often quoted remark from the introduction is that, according to Carson, “all such schemes [directed towards narrowing the bandwidth] are believed to involve a fundamental fallacy”. Among others, Gabor [10] takes this statement as a first step towards the understanding that bandwidth limitation sets a fundamental limit to the possible information transfer rate of a system.

The Significance of Noise

We have seen that the work towards a general theory of communication had two major breakthroughs where several investigators made similar but independent discoveries. The first one came in the 1920s by the discovery of the relation between bandwidth, time and information rate. The second one came 20 years later. An important difference between the theories published during these two stages is that in the 1920s the concept of noise was completely lacking.

Why did it take twenty years to fill the gap between Hartley's law and Shannon's formula? The only necessary step was to substitute $1 + C/N$ for m in (4). Why, all of a sudden, did three or more people independently “see the

¹²⁾ A paper [21] with this title by Colpitts and Blackwell was published in three instalments in Journal of the American Institute of Electrical Engineers in 1921, giving a comprehensive overview both of the history of the subject and the state-of-the-art around 1920.

¹³⁾ This was an intuitively appealing idea at the time, but sounds almost absurd today, when the bandwidth expansion of FM is a well-known fact.

light” almost at the same time? Why did neither Nyquist, nor Hartley or Küpfmüller realize that noise, or more precisely the signal-to-noise ratio play as significant a role for the information transfer capacity of a system as does the bandwidth?

One answer might be that they lacked the necessary mathematical tools for an adequate description of noise. At this time Wiener had just completed a series of papers on Brownian motion (1920–24) which would become a major contribution to what was later known as stochastic processes, the standard models for description of noise and other unpredictable signals, and one of Shannon’s favourite tools. These ideas, based on probabilistic concepts, were, however, far from mature to be used by even the most sophisticated electrical engineers of the time¹⁴⁾. Against this explanation, it may be argued that when Shannon’s formula was discovered, only two¹⁵⁾ of the independent researchers used a formal probabilistically based argument. The others based their reasoning on more common-sense reasoning, not resorting to other mathematical techniques than what were tools of the trade in the 1920s.

Another explanation could be that the problem of noise was rather new at the time. Noise is never a problem as long as it is sufficiently small compared to the signal amplitude. Therefore it usually arises in situations where a signal has been attenuated during transmission over a channel. As we have seen, such attenuation had been a problem from the early days of telegraphy. From the beginning, the problem of attenuation was not that noise then became troublesome, but rather that the signal disappeared altogether. (More precisely, it became too weak to activate the receiving apparatus in the case of telegraphy, or too weak to be heard by the human ear in case of telephony.) This situation changed radically around 1910, when the first practical *amplifiers* using vacuum tubes were devised. By use of these, long-distance telephony could for the first time be achieved, and ten years later the first commercial radio broadcasting could begin. But, alas, an electronic amplifier is not able to distinguish between signal and noise. So, as a by-product, interference, such as thermal noise¹⁶⁾, always present in both the transmission lines and the components of the amplifiers, would be amplified – and made audible – together with the signal. Early amplifiers were not able to amplify the signal very much, due to stability problems.

When the feed-back principle, patented by H.S. Black in 1927, came in use, gains in the order of hundreds or thousands became possible by cascading several amplifier stages. This made noise a limiting factor to transmission systems, important to control, and by the 1930s signal-to-noise ratio had become a common term among communications engineers.

Although the researchers of the 1920s were aware of the practical problem represented by noise and interference, it seems that they did not regard it as a *fundamental* property of the transmission system, but rather as one of the many imperfections of practical systems that should be disregarded when searching for principal limits of what could be accomplished.

Thus, the fact that noise had just begun to play an active role in communications systems, might partly explain why it was not given sufficient attention as a limitation to transmission capacity. However, when one looks at the reasoning used by both Tuller and Clavier (and to some degree Shannon), one will find that their arguments are inspired by two practical ideas, both invented during the 1930s, namely *frequency modulation (FM)* and *pulse code modulation (PCM)*.

Two Important Inventions

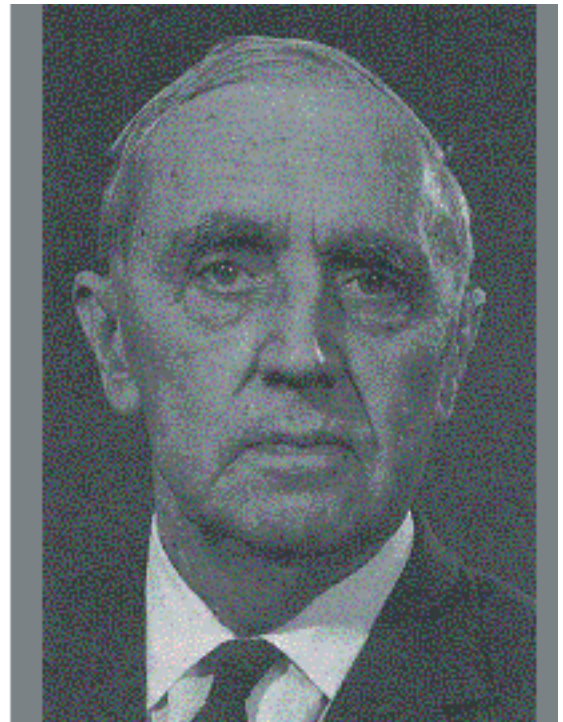
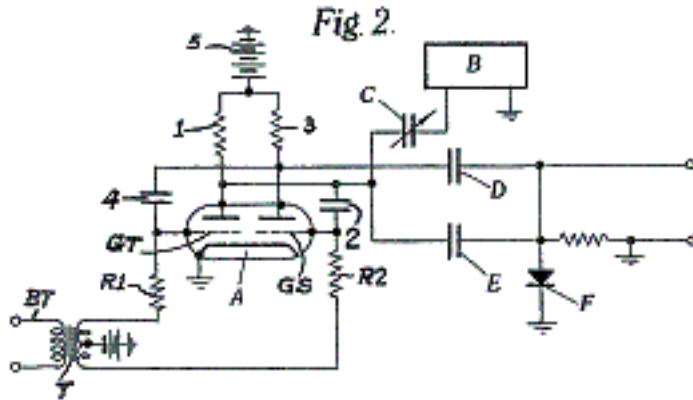
When reading textbooks and taking university courses in engineering, one may get the idea that new products are based on the results of engineering science, which again rely on a thorough understanding of more basic sciences such as physics and chemistry, which finally lean on mathematics as the most basic of all exact knowledge. One can also get the impression that the development in these fields follows the same pattern: technology must wait for physics to explain new phenomena by mathematics made ready for the purpose in advance. Reality is quite different. Time and time again inventors with only coarse knowledge of the physical phenomena they exploit, have come up with ingenious problem solutions. Similarly, engineers and physicists have discovered several mathematical results, which they have not been able to prove satisfactorily, leaving it to the established mathematicians to “tie up the ends” and provide the necessary comprehensive theory afterwards.

With this in mind, it is interesting to note that the understanding of noise in a theory of channel capacity had to wait for two practical inventions. These inventions illustrated how signal-to-noise

¹⁴⁾ One of the first papers with a rudimentary mathematical treatment of noise was published by Carson in 1925 [12]. It should also be mentioned that Harry Nyquist derived a mathematical model of thermal noise in 1928 [13]. This model was, however, derived without use of probabilistic methods.

¹⁵⁾ Shannon and Wiener, of course.

¹⁶⁾ Not to mention the “shot noise” generated by travelling electrons in the tubes themselves.



ratio (SNR) and bandwidth of a transmission system could actually be traded one against the other.

We have already mentioned Carson's 1922 paper, where he showed that frequency modulation (FM) would result in a signal bandwidth considerably larger than what would result by SSB, or even traditional AM. This is undeniably true, and it was therefore natural that most researchers also accepted Carson's rejection in the same article that FM would be more robust with respect to noise. Among the few who continued working seriously with FM was Edmund Armstrong. After several years of experimentation, and after introducing an *amplitude limiter* in the receiver, he was able to demonstrate that it was possible to significantly increase the SNR of a radio communication system by using FM at the cost of expanded bandwidth¹⁷⁾ [13].

What Armstrong's results showed was that a trade-off between bandwidth and SNR could in principle be possible. The next step, to realize that the information transfer capacity of a system depended both on bandwidth and SNR took some time, and needed another invention.

PCM – Pulse Code Modulation – consists of the sampling and quantizing of a continuous waveform. The use of sampling in telephone transmission had been suggested as early as 1903

by W.M. Miner.¹⁸⁾ Miner's motivation was to enhance the capacity of transmission lines by time division multiplex, as had already been done for telegraphy (see above). Miner's concept contained no form of quantizing and should not be considered a PCM system. This important addition was first introduced by A.H. Reeves in 1937.¹⁹⁾ Reeves realized that his system would need more bandwidth than traditional modulation methods. His rationale was the same as Armstrong's: the combat of noise. Reeves' radical insight was that by inserting repeaters at suitable intervals along the transmission line, no additional noise would be added during transmission together with the quantizing noise introduced by the encoding (modulation) process at the transmitter. The quantizing noise could be made arbitrarily small by using a sufficiently high number of quantization levels.

Implicit in Reeves' patent lies two important principles:

1. An analog signal, such as speech, can be represented with arbitrary accuracy by use of sufficiently frequent sampling, and by quantizing each sample to one of a sufficiently large number of predefined levels.
2. Each quantized sample can be transmitted on a channel with arbitrarily small probability of error, provided the SNR is sufficiently large.

Alec Harley Reeves (1902 – 1971) with two figures from his PCM patent. Although aware that his invention was far ahead of what was possible with the technology of his time, the patent included several circuit solutions to some of the involved functionalities

¹⁷⁾ Carson immediately accepted this as a fact, and very soon afterwards presented a paper together with C. Fry [15] showing mathematically how this mechanism worked, and that this was due to the amplitude limiter not included in the early proposals refuted by him in 1922.

¹⁸⁾ U.S. Patent 745,734.

¹⁹⁾ French Patent 852,183.

A result from this, not explicitly stated by Reeves, is that on a noise-free channel, an infinite amount of information can be transmitted in an arbitrarily small bandwidth. This is in sharp contrast to the results of the 1920s, and should be considered as a major reason why “Shannon’s formula” was discovered by so many just at a time when PCM was starting to become well-known.

Concluding Remarks

In this paper we have not written much about Shannon’s general information theory. On the other hand, we have tried to show how the ideas leading up to “Shannon’s formula” gradually emerged, as practical technological inventions made such ideas relevant. We have also seen that after the end of World War II, the subject was sufficiently mature, so that several independent researchers could complete what had been only partially explained in the 1920s.

On this background, one might be led to conclude that Shannon’s work was only one among the others, and, by a stroke of luck, he was the first one to publish his results. To avoid such a misunderstanding, we will briefly indicate how Shannon’s work is clearly superior to the others. First, we should make a distinction between the works of Shannon and Wiener ([1]–[3]) and the others ([5] [16] [17]). Both Shannon and Wiener delivered a general information measure based on the probabilistic behaviour of information sources, which they both designate by *entropy* due to the likeness with similar expressions in statistical mechanics. Shannon, furthermore, uses this concept in his general definition of channel capacity:

$$C = \max[H(x) - H_y(x)].$$

This expression can be interpreted as the maximum of the difference of the uncertainty about the message before and after reception. The result is given in bit/second and gives an upper bound of how much information can be transmitted *without error* on a channel. The most astonishing with Shannon’s result, which was not even hinted at by Wiener, was perhaps not so much the quantitative expression as the fact that completely error-free information exchange was possible at *any* channel, as long as the rate was below a certain value (the channel capacity).

The entropy concept is absent in all the presentations of the other group, which deal explicitly with a channel with additive noise. All reasoning is based on this special case of a transmission channel. The genius of Shannon was to see that the role of noise (or any other disturbances, being additive or affecting the signal in any other way) was to introduce an element of *uncertainty*

in the transmission of symbols from source to destination. This uncertainty is adequately modelled by a probability distribution. This understanding was shared by Wiener, but his attention was turned in other directions than Shannon’s. According to some, Wiener “under the misapprehension that he already knew what Shannon had done, never actually found out” [4].

References

- 1 Shannon, C E. A Mathematical Theory of Communication. *Bell Syst. Techn. J.*, 27, 379–423, 623–656, 1948.
- 2 Wiener, N. *Cybernetics: or Control and Communication in the Animal and the Machine*. Cambridge, MA, MIT Press, 1948.
- 3 Shannon, C E. Communication in the Presence of Noise. In: *Proc. IRE*, 37, 10–21, 1949.
- 4 Pierce, J R. The Early Days of Information Theory. *IEEE Trans. on Information Theory*, IT-19 (1), 1973.
- 5 Tuller, W G. Theoretical Limitations on the Rate of Information. *Proc. IRE*, 37 (5), 468–78, 1949.
- 6 Carson, J R. Notes on the Theory of Modulation. *Proc. IRE*, 10, 57, 1922.
- 7 Nyquist, H. Certain factors affecting telegraph speed. *Bell Syst. Techn. J.*, 3, 324–352, 1924.
- 8 Nyquist, H. Certain topics in telegraph transmission theory. *AIEE Trans.*, 47, 617–644, 1928.
- 9 Hartley, R V L. Transmission of information. *Bell Syst. Techn. J.*, 7, 535–563, 1928.
- 10 Küpfmüller, K. Über Einschwingvorgänge in Wellen Filtern. *Elektrische Nachrichtentechnik*, 1, 141–152, 1924.
- 11 Gabor, D. Theory of communication. *J. IEE*, 93 (3), 429–457, 1946.
- 12 Carson, J R. Selective Circuits and Static Interference. *Bell Syst. Techn. J.*, 4, 265, 1925.
- 13 Nyquist, H. Thermal Agitation of Electric Charge in Conductors. *Phys. Rev.*, 32, 1928.
- 14 Armstrong, E H. A Method of Reducing Disturbances in Radio Signaling by a System of Frequency-Modulation. *Proc. IRE*, 24, 689–740, 1936.

- 15 Carson, J R, Fry, T C. Variable Frequency Circuit Theory with Application to the Theory of Frequency-Modulation. *Bell Syst. Tech. J.*, 16, 513–540, 1937.
- 16 Clavier, A G. Evaluation of transmission efficiency according to Hartley's expression of information content. *Elec. Commun. : ITT Tech. J.*, 25, 414–420, 1948.
- 17 Laplume, J. Sur le nombre de signaux discernables en présence du bruit erratique dans un système de transmission à bande passante limitée. *Comp. Rend. Adac. Sci. Paris*, 226, 1348–1349, 1948.
- 18 Carson, J. The statistical energy-frequency system spectrum of random disturbances. *Bell Syst. Tech. J.*, 10, July, 374–381, 1931.
- 19 Oslin, G P. *The Story of Telecommunications*. Georiga, Macon, 1992.
- 20 Blanchard, J. The History of Electrical Resonance. *Bell Syst. Techn. J.*, 23, 415–433, 1944.
- 21 Colpitts, Blackwell. Carrier Wave Telephony and Telegraphy. *J. AIEE*, April 1921.
- 22 Bray, J. *The Communications Miracle*. New York, Plenum Press, 1995.
- 23 Hagemeyer, F W. *Die Entstehung von Informationskonzepten in der Nachrichtentechnik : eine Fallstudie zur Theoriebildung in der Technik in Industrie- und Kriegsforschung*. Berlin, Freie Universität Berlin, 1979. (PhD dissertation.)

Statistical Communication Theory 1948 – 1949

NIC. KNUDTZON ¹⁾



Dr. Nic. Knudtzon (80) obtained his Engineering degree from the Technical University of Norway, Trondheim in 1947 and his Doctor's degree from the Technical University in Delft, the Netherlands in 1957. 1948–1949 he was with the Research Laboratory of Electronics, Massachusetts Institute of Technology, working with information theory and experiments. 1950–1955 he was with the Norwegian Defence Research Establishment, Bergen, working on the development of microwave radio links; and from 1955 to 1967 he was Head of the Communications Division at Shape Technical Center in The Hague, Netherlands, where his efforts went into the planning of military telecommunications networks and systems in Western Europe. From 1968 to 1992 he was Director of Research at the Norwegian Telecommunications Administration, working on the planning of future telecommunications systems, networks and services. Dr. Knudtzon has been member of government commissions and various committees, including the Norwegian Research Council, the National Council for Research Libraries, the International Telecommunications Union, EURESCOM, etc.

Introduction

I thank you for the invitation to once again enter this distinguished rostrum here at Telenor's R&D Department!

My presentation will consist of three parts:

1) *The general state of affairs, 1948–1949:*

In order for you to properly understand and appreciate the topic I have been given, which dates back more than 50 years, I initially need to place *you* in the world of those times.

2) *Statistical communication theory, 1948–1949:*

Consistent with the original terminology, I shall use statistical communication theory as a common term for Claude Shannon's information theory and Norbert Wiener's mathematical theory of statistically optimal networks. This part of the presentation is an overview of what I saw, heard, and learned about the topic during my stay at the MIT Research Laboratory of Electronics from February 1948 until July 1949. I had the good fortune – as the only Norwegian – to work in this most inspiring environment during those pioneering years. Furthermore I shall give a very condensed reprise of the three papers I presented on the topic at "Studiemøtet i radioteknikk og elektroakustikk" (Symposium of Radio Technology and Electro-acoustics) at Farris Bad in 1950.

3) *Reflections:* Finally, I will conclude with some reflections on the developments – positive and negative – since those days.

The General State of Affairs, 1948 – 1949

Let us now place ourselves in 1948–1949 and list some notable events during this period:

- The relationship between the western world and The Soviet Union is dominated by the cold war.
- The new German states of Western Germany and Eastern Germany are established.
- Harry Truman unexpectedly wins the presidential election in the USA.

- The Marshall Plan is initiated.
- NATO is established, with Norway as a member.
- Mao Tse-Tung proclaims the People's Republic of China.
- The state of Israel is proclaimed.
- Mahatma Gandhi is murdered by a Hindu fanatic in Delhi.
- Long-playing records are introduced in the USA.
- The number of television receivers is reaching 750,000 in the USA.
- A committee is appointed to assess television in Norway.
- The transistor is invented.
- Claude E. Shannon publishes "A Mathematical Theory of Communication".
- Norbert Wiener publishes the book "Cybernetics".

The State of Science and Research in Norway, 1948

In 1941 we were 30 students who, based on our outstanding results from the matriculation exam, were admitted to The Faculty of Electrical Engineering at The Norwegian Institute of Technology (NTH): 20 to Power Electrical Engineering, and 10 to Electronics. In this wartime period the Faculty had three professors, all in Power Electrical Engineering. Their expertise was also made available to the Electronics students, although the frequency range was limited to 50 Hz. However, in the second half of our studies, laboratory engineer *Reno Berg* introduced us to frequencies above 50 Hz.

Our most memorable experience was the six weeks during the winter of 1946 when *Helmer Dahl* (38 at the time) and *Matz Jenssen* (36) – based on their achievements and experience

¹⁾ *The paper is a transcript of a presentation given by the author at the Telenor seminar "From Shannon's information theory to today's information society: Claude Shannon (1916–2001) In Memoriam" on August 9, 2001. The paper was presented in Norwegian and has been translated by Geir E. Øien.*

from UK laboratories during World War II – gave us unforgettable inspiration by providing insights into the enormous progress made in our scientific discipline. The frequencies reached into the GHz range!

My Personal Situation

Please allow me to say a few words about my personal situation, which led me to the USA and to close contact with the pioneers of statistical communication theory.

During his stay at NTH in early 1946, Helmer Dahl gave a lecture on a topic which, in his own words, “*had nothing to do with the curriculum*”: Thermal noise as the fundamentally limiting factor of communication systems. This caught my attention to such a degree that the choice of topics both for my main thesis and my doctoral thesis was made there and then! Since NTH had no expertise in this area, I wrote to Helmer Dahl, who at the time was establishing the so-called Department of Radar – in fact their research activities came to be primarily associated with microwave radio links – at The Norwegian Defence Research Institute (FFI), and asked whether it would be possible for me to do my main thesis under his supervision.

The topic of the thesis was remarkable for Norway at the time: “*Give an overview of molecular noise (fluctuation noise), and compute the signal-to-noise ratio associated with pulse-width modulation and frequency modulation for ultra-shortwave transmission.*” This was the first step in a long-term strategy to determine the choice of modulation scheme for radio links developed by FFI. The two alternatives were frequency modulation, which is compatible with conventional carrier frequency telephony, and pulse modulation with time-shared channels. Hence, in 1947 I was already told to assess fluctuation noise and compute signal-to-noise ratios in a *digital* system, which could not be efficiently realized in those pre-transistor days, but which was possible to analyse mathematically.

My studies were successful, and FFI sent me to Massachusetts Institute of Technology (MIT) for a period of 5 months, which MIT generously offered to extend by one year. Travelling to the USA in those days was not a trivial affair; I had to take an oath – with my hand on the Bible – as I received my government official visa. Subsequently I travelled by an ocean liner over the Northern Atlantic Ocean, in a hurricane.

USA 1948 – 1949

The transition from a war-ridden and run-down Norway, whose reconstruction had not yet properly started, to the dynamic USA, was overwhelming. As a small illustration, I may mention

that both NTH and MIT were building new libraries at the time: At NTH they dug with spades and transported by horse-carriages; at MIT I saw a bulldozer in action for the first time.

I came to a flourishing and proud USA, the winner of World War II, now acknowledged as the world’s leading superpower.

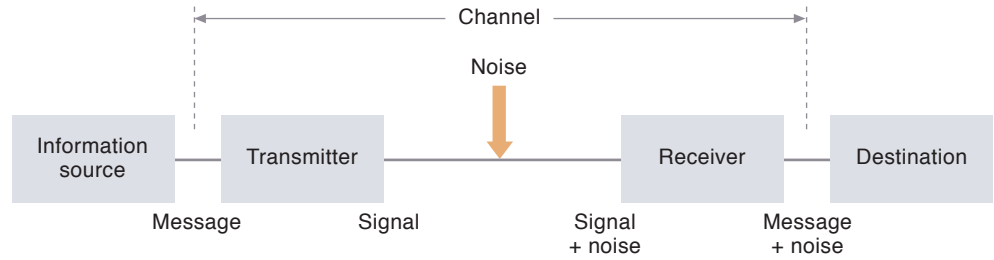
At MIT I worked at the Research Laboratory of Electronics, in the temporary buildings of the Radiation Laboratory, which had been a center for the US research on radar during World War II.

The annual updating within our scientific disciplines took place at the IRE (Institute of Radio Engineers) National Convention in New York, and my first meeting with this mustering of theory and practice, shortly after my arrival in 1948, was a great experience. In my diary I commented: “*Good*” (!) – particularly on a March 24 session, which included contributions from Claude Shannon and Norbert Wiener, whom I then saw and heard for the first time. Claude Shannon’s contribution was “*Communications in the presence of noise*” [1].

Of the many other impressions, I would here particularly like to emphasize the immediate, friendly American way of communicating and the sense of being taken seriously – no matter your age – when you accepted a challenge. Claude Shannon serves as a typical example in his relations with me. When he, on May 17, 1948, visited the group of 5–6 researchers to which I belonged at MIT, he invited me to visit Bell Telephone Laboratories. During the whole of July 7 he was my cicerone, both at the old location in West Street, and at Murray Hill. There I also met other well-known scientists, among them H. Nyquist and S.O. Rice, whose articles “*Mathematical Analysis of Random Noise*” [2] and “*Statistical Properties of a Sine Wave plus Random Noise*” [3] have given me the background for much of my work. On August 10, I personally received from Claude Shannon a copy of the preprint of his pioneering paper “*A Mathematical Theory of Communication*”, and with this treasure I immediately initiated a seminar series in our research group at MIT.

Claude Elwood Shannon (1916 – 2001) acquired his B.Sc. degrees in electrical engineering and mathematics at Michigan University in 1936, and his M.Sc. degree in electrical engineering as well as his Ph.D. degree in mathematics at MIT in 1940. He became a National Research Fellow in 1941. Subsequently he worked at Bell Telephone Laboratories and had an affiliation with

Figure 1 General Communication System

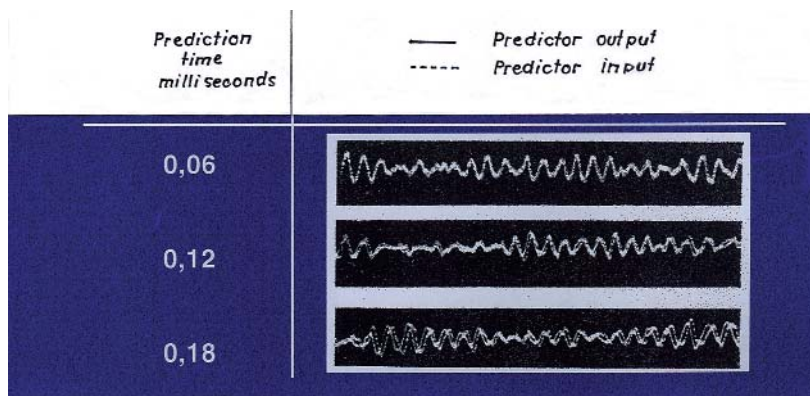


this institution until 1972. He was a professor at MIT from 1956 until 1978, when he became *Professor Emeritus*. Apart from his monumental contributions in statistical communication theory he also wrote papers in several other areas, while at the same time serving as an advocate of moderation regarding publishing for the sake of publishing itself. The most spectacular I have heard about him, is his passion for riding a unicycle while juggling. Many have offered the opinion that Shannon was worthy of a Nobel Prize, but it seems like he fell outside the six Nobel categories. My meetings with Claude Shannon took place when he was 32–33 years old and at the top of his most productive period, and I remember him as a pleasant and straightforward man.

Norbert Wiener (1894 – 1964) was in 1948 a world-famous 52-year-old mathematician who made his daily rounds in the MIT corridors and laboratories, a bit remote and with a particular sense for absent-mindedness. Due to his interest in statistical communication theory he visited our research group quite often, also observing some of my measurements of the statistical properties of noise. I sporadically followed some of his lectures in mathematics, which could be quite peculiar – particularly when he became addicted to his own computations at the blackboard, giving us associations to Maxwell’s Demon himself.

Norbert Wiener was affiliated with MIT since 1919, and was a professor there when Claude Shannon acquired his Ph.D. in mathematics.

Figure 2 Example of prediction of filtered fluctuation noise



Statistical Communication Theory

The Development

In 1928, R.V.L. Hartley at Bell Telephone Laboratories concluded in his “*Transmission of Information*” [5], based on physical (as opposed to psychological) considerations, that the amount of information transmitted in a noiseless system is proportional to the product of bandwidth and transmission time.

In 1946 this result was developed further by D. Gabor at British Thomson-Houston Co. in “*Theory of Communication*” [6], but he, too, confined himself to a system without noise.

At a meeting of the IRE in New York on December 12, 1947, A.G. Clavier of ITT’s Federal Telecommunication Laboratories presented “*Evaluation of Transmission Efficiency According to Hartley’s Expression of Information Content*” [7]. This work contained an analysis of the transmission efficiency of frequency modulation and various kinds of pulse modulation, based on Hartley’s definition of information content, but for a system *with* noise.

At the same IRE meeting in New York, Claude Shannon presented “*Communications in the presence of noise*” [1], which is identical to his previously mentioned presentation made at IRE’s National Convention in 1948. He took as a starting point his figure of a “General Communication System” (cf. Figure 1), used a geometrical viewpoint when solving the problem, and concluded with his formula for the optimal transmission capacity,

$$C = W \log_2 (P + N) / N,$$

for a communication channel with bandwidth W , average transmit power P , and thermal (Gaussian and white) noise power N . Furthermore, he presented figures showing how practical systems compared in quantitative performance to the optimal performance theoretically attainable. I discuss this paper in depth in [11] and [12].

Then, in July 1948, Claude Shannon's pioneering paper [4] was published. This paper, too, is discussed in [11] and [12]. Claude Shannon, in a footnote (p. 626 in [4]), expresses the view that "Communication theory is heavily indebted to Wiener for much of its basic philosophy and theory ...". He also states (p. 627) that "We may also here refer to Wiener's forthcoming book "Cybernetics" dealing with the general problems of communication and control." Without going further into the matter, it should also be noted that Claude Shannon worked on cryptography during World War II. This report was not declassified until 1948, when it was published as "Communication Theory of Secrecy Systems" [8].

During World War II Norbert Wiener developed "The Extrapolation, Interpolation, and Smoothing of Stationary Time Series" (National Defence Research Council, Section D2 Report, Feb. 1942), dealing with control of gunfire aimed at targets in motion. This report was also classified, and in addition rather heavy reading, so it stayed fairly unknown before it became the basis for an MIT course from 1947. I belonged to the second class following this course. Norbert Wiener's starting point is an analysis of statistically stationary time series, a topic I discuss in [11] and [13]. For illustration, I have chosen to reproduce here – as Figure 2 – a simplified Figure 5 from [13], showing the result of using a statistically optimal predictor for filtered fluctuation noise – built at MIT by my good friend Charles A. Stutt as part of his Ph.D. work.

In 1948 Norbert Wiener published his book "Cybernetics", about control through communication between, and feedback of the processes in living organisms, in machines, and in society [9]. The title (derived from the Greek word for *steersman*) is not only contextually meaningful, it has also been included in the vocabulary of many languages. In Chapter III, "Time Series, Information, and Communication" (pp. 74-112), Norbert Wiener gives the following "certification": "The relevant general theory has been presented in a very satisfactory form by Dr. C. Shannon". Claude Shannon, in his book review of "Cybernetics" [10], writes: "Communication engineers have a charter right and responsibility in several of the roots of this broad field and will find Wiener's treatment interesting reading, filled with stimulating and occasionally controversial ideas. – Professor Wiener, who has contributed much to communication theory, is to be congratulated for writing an excellent introduction to a new and challenging branch of science."

The "Bit" as a Unit of Information

The information unit bit for "binary digit" seems to be officially introduced in print for the first time in [4], and was "suggested by J.W. Tukey" according to Claude Shannon. John Wilder Tukey (1915 – 2000) was a well-known statistician, a contemporary of Shannon from Bell Telephone Laboratories (birth and death one year before Shannon's), and at the same time professor at Princeton University. He also introduced the term "software", and made important contributions to the development of Fast Fourier Transform (FFT) algorithms.

My 1950 Presentations of Statistical Information Theory

My journal papers [11], [12], and [13] were the first presentations on statistical information theory in Norway.²⁾ I would like to point out one thing regarding these papers: The parts dealing with *information theory* are heavily influenced by Claude Shannon's own way of presenting his theory. Please note how many fundamental results can be derived using simple explanations. Following Shannon's publications many others have come up with articles and books, many of which try to glorify the subject matter by using needlessly complicated descriptions. My advice has always been: "Read Shannon in original!"

Reflections

On Mathematical Solutions

It may be worth pointing out that the methods of research have changed considerably as computers have become widely available and obtained greater computational power.

Wiener's solution for statistically optimal filters, as developed during World War II, was based on the choice of *minimum mean squared error* as an optimality criterion. This led to the Wiener-Hopf equations, which were solvable analytically. Without further comparison intended, I was one of those facing a similar situation in my doctoral work. We were dependent on mastering the analytical solutions; otherwise our research might fail completely, even after years of work. Today, computers are available for obtaining numerical solutions. However, it might be added that uncritical use of computers can also lead to a loss of judgement, intuition, and physical understanding.

Good or Bad?

The development of telecommunications is steadily accelerating, particularly after the advent of transistors, satellites, and optical fibres

²⁾ Translator's note: Translated versions of [11], [12] and [13] are found in this volume of *Teletronikk*.

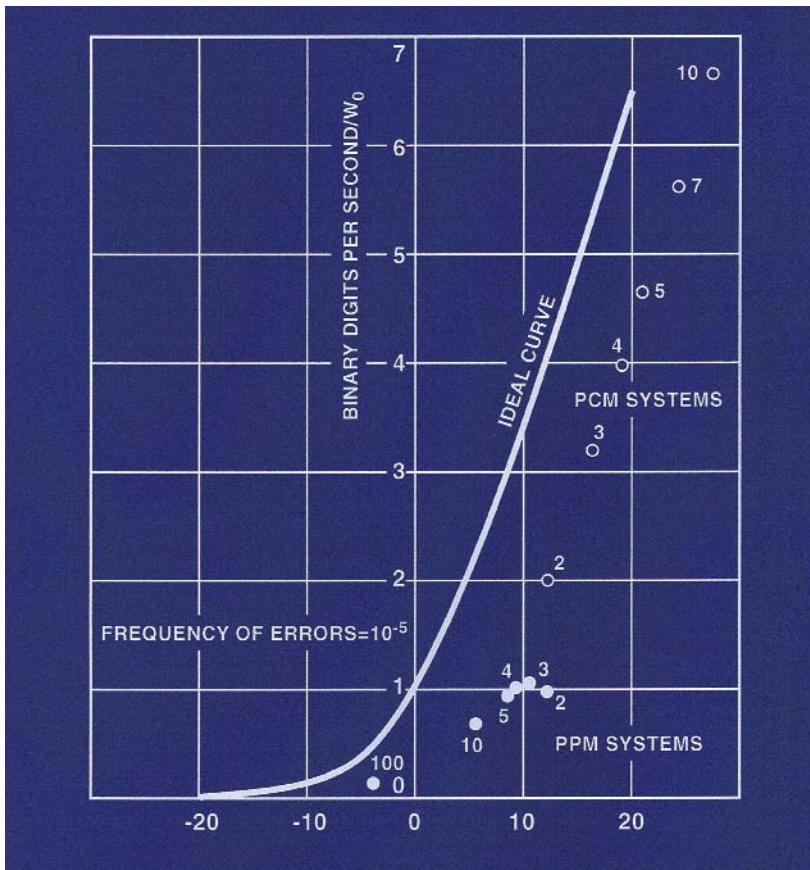


Figure 3 Comparison of PCM and PPM with ideal performance

in practical systems. The number of services available to the users is already overwhelming, and seems set to increase further.

As with all technology, telecommunications are in many cases a good servant, but it can also be a bad master. Personally, I particularly dislike three things:

- The mindless *over-use*. I am increasingly surprised to see that people have so much time and money to spend on telecommunication services.
- The *lack of quality in the content* that is transmitted, be it the shallowness of many mobile phone conversations, bad television programs, or the lack of quality control concerning what becomes available on the Internet.
- The *distortion of our language* found in e-mail.

However, regardless of

- type of communication link – copper wires, cable, fibres, or radio communication, direct or via satellite, fixed or mobile;
- user equipment – telegraph, telephone, facsimile, television, or computers;

Claude Shannon's fundamental capacity formula defines the optimal channel use, against which the various technical solutions can be measured;

see Figure 3. Shannon gave the system planners the basis for quantitative assessments.

References

- 1 Shannon, C E. Communications in the presence of noise. *Proc. of the IRE*, 10–21, 1949.
- 2 Nyquist, H. Mathematical analysis of random noise. *Bell Syst. Technical Journal*, XXIII, 282–332, 1944; and XXIV, 46–156, 1945.
- 3 Rice, S O. Statistical properties of a sine wave plus random noise. *Bell Syst. Technical Journal*, XXVII, 109–157, 1948.
- 4 Shannon, C E. A mathematical theory of communication. *Bell Syst. Technical Journal*, XXVII, 623–656, 1948; and XXVIII, 623–715, 1948. (Later reprinted in: Shannon, C E, Weaver, W. *The Mathematical Theory of Communication*. The University of Illinois Press: Urbana, 1949.)
- 5 Hartley, R V L. Transmission of information. *Bell Syst. Technical Journal*, VII, 535–573, 1928.
- 6 Gabor, D. Theory of communication. *Journal of IEE*, 93, 439–457, 1946.
- 7 Glavier, A G. Evaluation of transmission efficiency according to Hartley's expression of information content. *Electrical Communication*, 25, 414–420, 1948.
- 8 Shannon, C E. Communication theory of secrecy systems. *Bell Syst. Technical Journal*, XXVIII, 656–715, 1949.
- 9 Wiener, N. *Cybernetics or The Control and Communication in the Animal and the Machine*. New York, The Technology Press, John Wiley, 1948.
- 10 Shannon, C E. Cybernetics, or Control and Communication in the Animal and the Machine, by Norbert Wiener (book review). *Proc. of the IRE*, 1305, 1949.
- 11 Knudtzon, N. Statistisk kommunikasjonsteori. En kortfattet oversikt over problemstillingen. *Teknisk Ukeblad*, 16. nov. 1950, 883–887.
- 12 Knudtzon, N. Informasjonsteori. *Elektroteknisk Tidsskrift*, 30, 373–380, 1950.
- 13 Knudtzon, N. Statistisk optimale nettverk. *Elektroteknisk Tidsskrift*, 32, 413–416, 1950.

The True Channel Capacity of Pair Cables With Respect to Near End Crosstalk

NILS HOLTE



Nils Holte (55) is professor in Telecommunications at the Norwegian University of Science and Technology (NTNU). He received his Siv.Ing. degree in 1971 and his Dr.Ing. degree in 1976, both from NTNU. His main research interests are adaptive filters, crosstalk in pair cables, and digital communications, with special emphasis on OFDM and xDSL systems.

Nils.Holte@tele.ntnu.no

Previously, the channel capacity of pair cables that use two-way transmission has been calculated by means of a worst-case estimate of near end crosstalk (NEXT). This is a pessimistic estimate because near end crosstalk in frequency bands with some frequency separation is uncorrelated. In this paper the true channel capacity in a pair cable with respect to NEXT is calculated by means of a stochastic crosstalk model. It is shown how the mean and variance of the channel capacity can be calculated both by analytical methods and by simulation. The entire probability distribution of the channel capacity is estimated by means of a Monte Carlo simulation over a large ensemble of cables. Hence, a true estimate of the 1 % confidence limit of the channel capacity with respect to NEXT can be calculated for a given type of cable. The results are compared with traditional estimates. It is shown for a typical example that the worst-case channel capacity is approximately 5 % higher than the traditional estimate and that the average channel capacity is approximately another 10 % higher than the true worst-case estimate.

1 Introduction

Different types of xDSL (Digital Subscriber Line) systems used in the existing pair cables of the public access network are some of the major alternatives for implementing fixed broadband access. Symmetrical systems like SHDSL [1] have equal transmission rates in both directions and use two-way transmission within the same frequency band. In this case the dominating noise mechanism is near end crosstalk (NEXT). The traditional method for calculating the channel capacity of pair cables with respect to crosstalk has been to use Shannon's channel capacity formula, where the noise power spectral density is based upon the 99 % confidence limit of the crosstalk power sum. This approach is used in textbooks like Starr, Cioffi and Silverman [2] and also in different xDSL standards, for instance the ITU standard for SHDSL [1]. This method gives a pessimistic estimate for NEXT, because NEXT has significant variations with frequency. The NEXT noise powers in two different frequency bands are uncorrelated for a frequency separation greater than approximately 100 – 200 kHz. This effect was first modelled in detail by Gibbs and Addie [3] who applied their model to single carrier baseband systems.

In the current paper, the basic principles of Gibbs and Addie are generalised and extended to the calculation of channel capacity. This gives a realistic estimate of the achievable bitrate for a multicarrier system using adaptive modulation. The calculations are based both upon analytical methods and a Monte Carlo simulation of random crosstalk coupling coefficients throughout the cables in a large ensemble of cables. By means of the Monte Carlo simulation the entire

probability distribution of the channel capacity is calculated. Hence, the 1 % confidence limit of the channel capacity for a given system is estimated directly from the probability distribution of the channel capacity instead of the indirect estimate based upon worst case NEXT (99 % confidence limit).

The new method will primarily be suitable for systems with two-way transmission, where NEXT is the dominating noise mechanism. The results show that the new method gives an increase in bitrate of approximately 5 % compared to the traditional method when applied to a system with bandwidth 400 kHz in a 10 pair cable (10 pair binder group).

This new approach may in principle be used also for far end crosstalk (FEXT). However, in cables with identical propagation constants in all pairs, there will be full correlation between FEXT at different frequencies, so that the new and the traditional method will give identical results.

The paper is organised as follows. First, a NEXT model is presented together with an analytical analysis of NEXT. Then the Monte Carlo simulation is explained, and it is shown how the channel capacity may be calculated. Estimates of average rate and worst-case rate are presented for a multicarrier system example, and the results are compared with results found by using traditional estimates of NEXT. To conclude, it is explained how this new method might contribute to an improvement of the maximum range of transmission systems for the SHDSL application, by using a multicarrier system instead of the standardised single carrier SHDSL system.

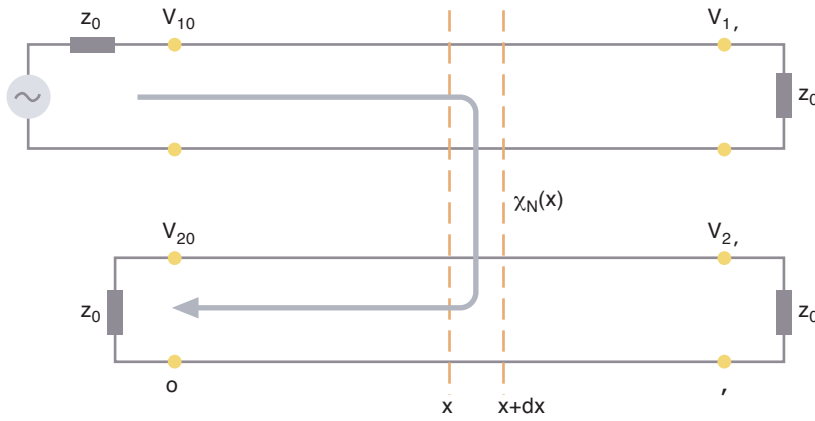


Figure 2.1 Near end crosstalk between two pairs

2 Near End Crosstalk Model

The near end crosstalk between two pairs in a multipair cable will consist of contributions as illustrated in Figure 2.1. It is assumed that both pairs are terminated by their characteristic impedance Z_0 at both ends of the cable.

The near end crosstalk transfer function for high frequencies ($f > 100$ kHz) between pair no. i and pair no. j of a cable is given by Klein [4]:

$$H_{Ni,j}(f) = \frac{V_{20}}{V_{10}} = j\beta_0 \int_0^l \mathcal{K}_{Ni,j}(x) \exp(-2\gamma x) dx. \quad (2.1)$$

$\beta_0 = \beta_0(f) = \omega\sqrt{LC}$ is the lossless phase constant of the cable in rad/km

$\gamma = \alpha + j\beta$ is the propagation constant of the cable

$\alpha = \alpha(f)$ is the attenuation constant of a pair in Neper/km

$\beta = \beta(f)$ is the phase constant of a pair in rad/km

$\mathcal{K}_{Ni,j}(x) = [C_{i,j}(x) / C + L_{i,j}(x) / L] / 2$ is the normalised NEXT coupling coefficient between pair i and j at position x

$C_{i,j}(x)$ is the mutual capacitance per unit length between pair i and j

$L_{i,j}(x)$ is the mutual inductance per unit length between pair i and j

C is the capacitance per unit length of the pairs

L is the inductance per unit length of the pairs

l is the cable length

According to Cravis and Crater [5], the coupling coefficients are stationary, Gaussian, random processes with correlation length less than a few meters. In comparison with actual values of α and β , this means that the coupling coefficients may be modelled as white noise processes with correlation functions:

$$R_{Ni,j}(\tau) = E[\kappa_{Ni,j}(x) \cdot \kappa_{Ni,j}(x + \tau)] = \kappa_{Ni,j} \delta(\tau). \quad (2.2)$$

The constants $\kappa_{Ni,j}$ will be different for different pair combinations and can be estimated from crosstalk measurements. Coupling coefficients in different pair combinations are statistically independent.

Most other authors [3, 5, 6] treat this coefficient as a random variable. Cravis and Crater [5] assume that $\kappa_{Ni,j}$ follows a gamma distribution, while Bradley [6] and Gibbs and Addie [3] assume a log normal distribution, but these assumptions are only approximations. For a given cable design the coefficient $\kappa_{Ni,j}$ is a deterministic function of the pair combinations, and it is mainly determined by the average distance between the two pairs and the difference in twisting periods between them [7].

A pair cable consisting of N pairs is used as an example. For simplicity, only one specific disturbed pair is taken into account, and this pair is denoted pair no. 1. The procedure will be the same for other disturbed pairs. Measurements for a 0.6 mm pair cross stranded cable using 10 pair binder groups have been averaged over individual pair combinations [8], and the result is shown in Table 2.1.

Pair combination	Average NEXT at 1 MHz
1-2	54.2 dB
1-3	55.7 dB
1-4	57.1 dB
1-5	57.9 dB
1-6	59.0 dB
1-7	59.0 dB
1-8	59.1 dB
1-9	59.3 dB
1-10	59.6 dB

Table 2.1 Average NEXT for different pair combinations in a 0.6 mm cross stranded pair cable

There are moderate differences in NEXT power sum between the pairs of a binder group (1.5 dB between max and min). The pair with the median of average NEXT power sum has been chosen as the disturbed pair (pair no. 1). Table 2.1 shows the average NEXT between pair no. 1 and all the other pairs in the binder group.

There will be a minor absolute error by using calculations for only one pair, but this error is probably less than 1 dB, and this approach will certainly be sufficient for a comparison of methods.

2.1 Average NEXT Transfer Functions

According to the model given by (2.1) and (2.2), the near end crosstalk transfer function will be a random variable with zero mean. The average NEXT power transfer function for one pair combination is given by:

$$P_{i,j}(f) = E \left[|H_{Ni,j}(f)|^2 \right] = E \left[\beta_0^2 \int_0^\infty \int_0^\infty \kappa_{Ni,j}(x) \cdot \kappa_{Ni,j}(y) \cdot \exp(-2\gamma x - 2\gamma^* y) \cdot dx \cdot dy \right]. \quad (2.3)$$

* denotes complex conjugate.

The upper integration limit has been increased

from 1 to infinity to simplify the calculations. This can be done because crosstalk coupling at the outer end of the cable does not contribute to NEXT for cable lengths of practical interest. Using (2.2) the result may be expressed:

$$P_{i,j}(f) = k_{Ni,j} \cdot \beta_0^2 \int_0^\infty \exp(-4\alpha x) \cdot dx = \frac{k_{Ni,j} \cdot \beta_0^2}{4\alpha}. \quad (2.4)$$

Above 100 kHz, the attenuation and phase constants may be approximated by:

$$\alpha = \alpha(f) = \alpha_{1M} \sqrt{F}, \quad (2.5)$$

$$\beta_0 = \beta = \beta(f) = \beta_{1M} F. \quad (2.6)$$

α_{1M} and β_{1M} are the attenuation and phase constants at 1 MHz.

F is the frequency in MHz.

Consequently the NEXT power transfer function is expressed:

$$P_{i,j}(f) = \frac{k_{Ni,j} \cdot \beta_{1M}^2}{4\alpha_{1M}} \cdot F^{1.5} = k_N \cdot F^{1.5}, \quad (2.7)$$

where k_N is a constant. This is the result found by Cravis and Crater [5], and it shows that average NEXT increases 15 dB/decade of frequency.

For a cable with many pairs, crosstalk from different pairs will add on a power basis. If the same type of system is used in all pairs of an N -pair cable, the effective crosstalk is given by the crosstalk power sum. Near end crosstalk power sum for pair no. 1 may be expressed:

$$|H_{Nps}(f)|^2 = \sum_{i=2}^N |H_{Ni,i}(f)|^2. \quad (2.8)$$

Average NEXT power sum will be:

$$P_{ps}(f) = E \left[|H_{Nps}(f)|^2 \right] = \sum_{i=2}^N E \left[|H_{Ni,i}(f)|^2 \right] = \frac{\beta_0^2}{4\alpha} \sum_{i=2}^N k_{Ni,i} = k_{Nps} \cdot F^{1.5}, \quad (2.9)$$

where k_{Nps} is a constant. NEXT power sum is also increasing 15 dB/decade with frequency.

2.2 The Covariance of NEXT

In order to study the correlation between NEXT at different frequencies it is convenient to investigate the covariance of the crosstalk power transfer function. The covariance of the NEXT crosstalk power transfer function at two frequencies f_1 and f_2 for one specific pair combination is defined by:

$$\text{cov} \left[|H_{Ni,j}(f_1)|^2, |H_{Ni,j}(f_2)|^2 \right] = E \left[|H_{Ni,j}(f_1)|^2 \cdot |H_{Ni,j}(f_2)|^2 \right] - P_{i,j}(f_1) \cdot P_{i,j}(f_2) \quad (2.10)$$

The first term may be expressed:

$$E \left[|H_{Ni,j}(f_1)|^2 \cdot |H_{Ni,j}(f_2)|^2 \right] = E \left[\beta_{01}^2 \cdot \beta_{02}^2 \int_0^\infty \int_0^\infty \int_0^\infty \int_0^\infty \kappa_{Ni,j}(x) \cdot \kappa_{Ni,j}(y) \cdot \kappa_{Ni,j}(z) \cdot \kappa_{Ni,j}(w) \cdot \exp(-2\gamma_1 x - 2\gamma_1^* y - 2\gamma_2 z - 2\gamma_2^* w) \cdot dx \cdot dy \cdot dz \cdot dw \right]. \quad (2.11)$$

The solution is in correspondence with Gibbs and Addie [3] given by:

$$E \left[|H_{Ni,j}(f_1)|^2 \cdot |H_{Ni,j}(f_2)|^2 \right] = \frac{k_{Ni,j}^2 \cdot \beta_{01}^2 \cdot \beta_{02}^2}{16} \left[\frac{1}{\alpha_1 \cdot \alpha_2} + \frac{4}{(\alpha_1 + \alpha_2)^2 + (\beta_1 - \beta_2)^2} + \frac{4}{(\alpha_1 + \alpha_2)^2 + (\beta_1 + \beta_2)^2} \right]. \quad (2.12)$$

It is common to use the assumption that $\alpha \ll \beta$ in order to simplify the results, and hence the last term may be neglected. However, this is only a fair approximation for the lowest frequencies. At 100 kHz the ratio between β and α is less than 10. Using this assumption, the covariance may be expressed:

$$\text{cov} \left[|H_{Ni,j}(f_1)|^2, |H_{Ni,j}(f_2)|^2 \right] = \frac{k_{Ni,j}^2 \cdot \beta_{01}^2 \cdot \beta_{02}^2}{4 \left[(\alpha_1 + \alpha_2)^2 + (\beta_1 - \beta_2)^2 \right]}. \quad (2.13)$$

The normalised covariance between the NEXT power transfer function for one pair combination at frequencies f_1 and f_2 is found by division with the average power sum at the two frequencies:

$$\rho = \frac{\text{cov}\left[|H_{Ni,j}(f_1)|^2, |H_{Ni,j}(f_2)|^2\right]}{P_{i,j}(f_1) \cdot P_{i,j}(f_2)} = \frac{4 \cdot \alpha_1 \cdot \alpha_2}{(\alpha_1 + \alpha_2)^2 + (\beta_1 - \beta_2)^2}. \quad (2.14)$$

The covariance of NEXT power sum is defined by:

$$\text{cov}\left[|H_{Nps}(f_1)|^2, |H_{Nps}(f_2)|^2\right] = E\left[\left(\sum_{i=2}^N |H_{Ni,i}(f_1)|^2\right) \cdot \left(\sum_{j=2}^N |H_{Ni,j}(f_2)|^2\right)\right] - P_{ps}(f_1) \cdot P_{ps}(f_2). \quad (2.15)$$

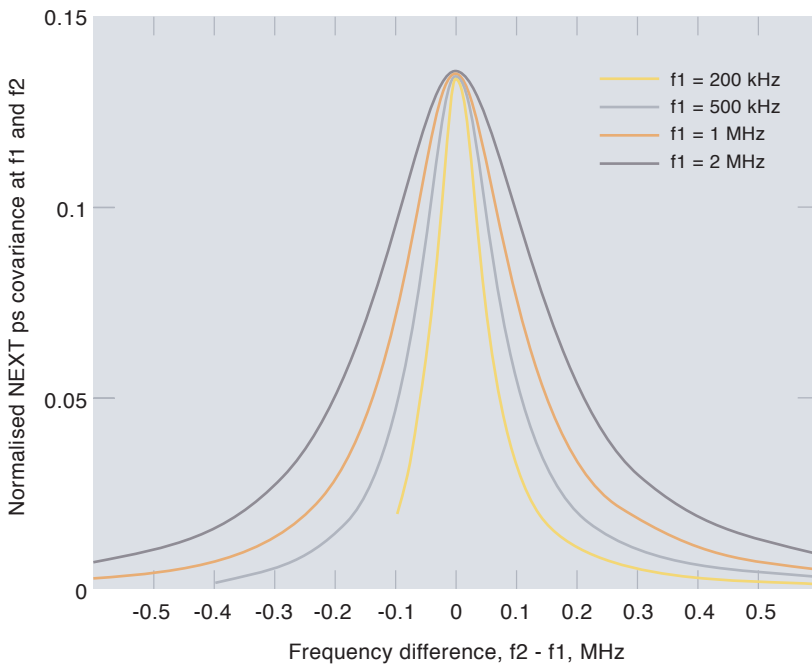
The result is expressed:

$$\text{cov}\left[|H_{Nps}(f_1)|^2, |H_{Nps}(f_2)|^2\right] = \frac{\beta_{01}^2 \cdot \beta_{02}^2}{4\left[(\alpha_1 + \alpha_2)^2 + (\beta_1 - \beta_2)^2\right]} \cdot \sum_{i=2}^N k_{Ni,i}^2. \quad (2.16)$$

The normalised covariance of NEXT power sum is:

$$\rho_{ps} = \frac{\text{cov}\left[|H_{Nps}(f_1)|^2, |H_{Nps}(f_2)|^2\right]}{P_{ps}(f_1) \cdot P_{ps}(f_2)} = \frac{4 \cdot \rho_{kN} \cdot \alpha_1 \cdot \alpha_2}{(\alpha_1 + \alpha_2)^2 + (\beta_1 - \beta_2)^2} = \rho \cdot \rho_{kN}. \quad (2.17)$$

Figure 2.2 The normalised covariance between NEXT power sum at two frequencies as a function of the frequency difference



where:

$$\rho_{kN} = \frac{\sum_{i=2}^N k_{Ni,i}^2}{\left[\sum_{i=2}^N k_{Ni,i}\right]^2}. \quad (2.18)$$

The approximations (2.5) and (2.6) are used for the frequencies F_1 and F_2 given in MHz. Setting $\varepsilon = \alpha_{1M} / \beta_{1M}$, the normalised covariance of NEXT power sum is expressed:

$$\rho_{ps} = \frac{4 \cdot \rho_{kN} \cdot \varepsilon^2 \cdot \sqrt{F_1} \cdot \sqrt{F_2}}{\varepsilon^2 \left(\sqrt{F_1} + \sqrt{F_2}\right)^2 + (F_1 - F_2)^2}. \quad (2.19)$$

For the 0.6 mm cable used as example, the coefficient ρ_{kN} is equal to 0.137 and $\varepsilon = 0.055$. The normalised covariance between NEXT power sum at f_1 and f_2 for this cable is shown in Figure 2.2.

The figure shows that near end crosstalk at two different frequencies with a large frequency separation is almost uncorrelated. The correlation is low for frequency differences larger than 100 kHz in the frequency range below 500 kHz.

2.3 Probability Distributions of NEXT

Crosstalk will vary statistically over an ensemble of different cables of the same type. The NEXT transfer function is a linear function of the coupling coefficients, and hence it will be a complex Gaussian variable with zero mean. In order to calculate the channel capacity of a cable it is desirable to know the probability distributions of NEXT. The real and imaginary parts of the NEXT transfer function for a single pair combination are given by:

$$r_{i,j}(f) = \text{Re}[H_{Ni,j}(f)] = j\beta_0 \int_0^\ell \mathcal{K}_{Ni,j}(x) \cdot \sin(2\beta x) \cdot \exp(-2\alpha x) dx, \quad (2.20)$$

$$q_{i,j}(f) = \text{Im}[H_{Ni,j}(f)] = j\beta_0 \int_0^\ell \mathcal{K}_{Ni,j}(x) \cdot \cos(2\beta x) \cdot \exp(-2\alpha x) dx. \quad (2.21)$$

A detailed calculation of the probability distribution of the crosstalk transfer function $H_{Ni,j}(f)$ is given in Appendix A. Under the assumption that $\alpha \ll \beta$, the second order moments of $r_{i,j}(f)$ and $q_{i,j}(f)$ are given by:

$$E[r_{i,j}^2(f)] = E[q_{i,j}^2(f)] = \frac{k_{Ni,j} \cdot \beta_0^2}{8\alpha} = \frac{P_{i,j}(f)}{2},$$

$$E[r_{i,j}(f) \cdot q_{i,j}(f)] = 0. \quad (2.22)$$

The power transfer function of NEXT at one frequency for a specific pair combination is denoted $z_{i,j}$, where $z_{i,j} = z_{i,j}(f) = r_{i,j}^2(f) + q_{i,j}^2(f)$. Because $r_{i,j}(f)$ and $q_{i,j}(f)$ are independent Gaussian variables with identical variances, the probability density of $z_{i,j}$ will in accordance with appendix A be given by:

$$P_{i,j}(z_{i,j}; f) = \frac{1}{P_{i,j}(f)} \cdot \exp\left(-\frac{z_{i,j}}{P_{i,j}(f)}\right), \quad z_{i,j} \geq 0. \quad (2.23)$$

NEXT power sum for pair no. 1 is denoted

$$z = z(f) = \sum_{i=2}^N z_{i,j}(f).$$

The probability distribution of NEXT power sum is given by:

$$p(z; f) = p_{1,2}(z_{1,2}; f) * p_{1,3}(z_{1,3}; f) * \dots * p_{1,N}(z_{1,N}; f), \quad (2.24)$$

where * denotes convolution.

For the general case where all pair combinations have different values of $k_{Ni,j}$, the result of the convolution is:

$$p(z; f) = \sum_{i=2}^N \left(\frac{[P_{1,i}(f)]^{N-3} \cdot \exp\left(-\frac{z}{P_{1,i}(f)}\right)}{\prod_{\substack{j=2 \\ j \neq i}}^N [P_{1,i}(f) - P_{1,j}(f)]} \right), \quad z \geq 0. \quad (2.25)$$

For the special case that all pair combinations have the same NEXT level, $P(f) = P_{1,i}(f)$, $i = 2, N$, the NEXT power sum will be gamma distributed with $N - 1$ degrees of freedom and probability density given by:

$$p(z; f) = \frac{1}{\Gamma(N-1)} \cdot \frac{z^{N-2}}{[P(f)]^{N-1}} \cdot \exp\left(-\frac{z}{P(f)}\right), \quad z \geq 0. \quad (2.26)$$

$\Gamma(x)$ is the gamma function. In practical cables the NEXT level usually varies between different pair combinations, so that (2.25) represents a more realistic case than (2.26).

2.4 Joint Probability Distribution of NEXT at Two Frequencies

In order to calculate estimates of the variance of for instance the channel capacity, it is necessary to know the joint probability distributions of NEXT at two different frequencies f_1 and f_2 . The following two-dimensional probability density function of NEXT power transfer function for a single pair combination is calculated in App-

endix B. The variables $z_{1,i,j}$ and $z_{2,i,j}$ are the NEXT power transfer function between pair i and j at frequencies f_1 and f_2 respectively.

$$p_{i,j}(z_{1,i,j}, z_{2,i,j}; f_1, f_2) = \frac{1}{P_{i,j}(f_1) \cdot P_{i,j}(f_2) \cdot (1 - \rho)} \cdot \exp\left(-\frac{z_{1,i,j}}{(1 - \rho) \cdot P_{i,j}(f_1)} - \frac{z_{2,i,j}}{(1 - \rho) \cdot P_{i,j}(f_2)}\right) \cdot I_0\left(\frac{2\sqrt{\rho}}{1 - \rho} \cdot \sqrt{\frac{z_{1,i,j} \cdot z_{2,i,j}}{P_{i,j}(f_1) \cdot P_{i,j}(f_2)}}\right). \quad (2.27)$$

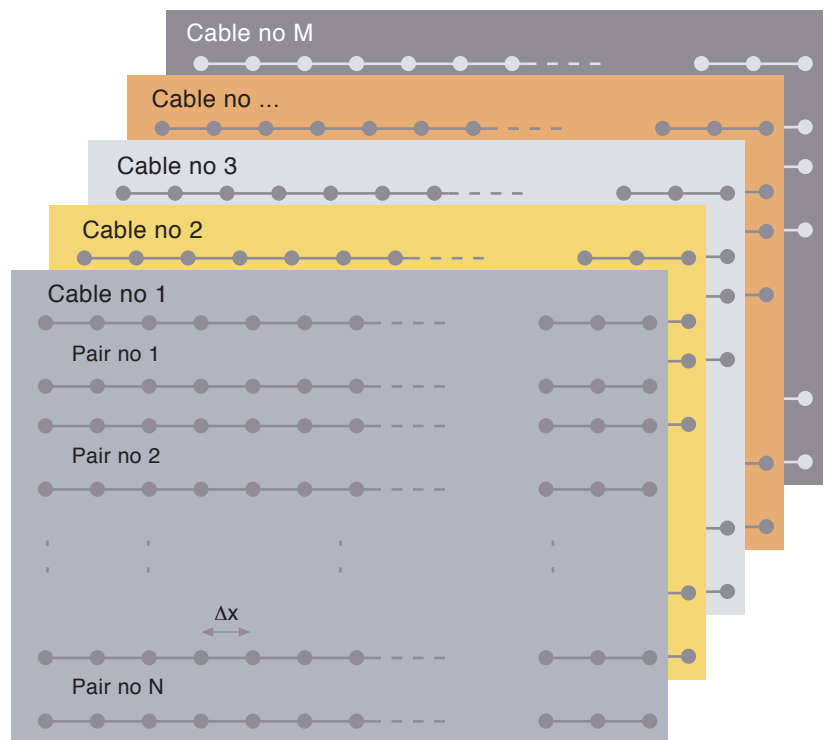
$I_0(x)$ is the modified Bessel function of order zero [14], and ρ is given by (2.14).

The two-dimensional probability density function of NEXT power sums $z_1 = z(f_1)$ and $z_2 = z(f_2)$ for pair no. 1 is given by:

$$p(z_1, z_2; f_1, f_2) = p_{1,2}(z_{1,1,2}, z_{2,1,2}; f_1, f_2) * p_{1,3}(z_{1,1,3}, z_{2,1,3}; f_1, f_2) * \dots * p_{1,N}(z_{1,1,N}, z_{2,1,N}; f_1, f_2). \quad (2.28)$$

In this equation * denotes two-dimensional convolution. Given that the two independent sets of parameters (x_1, y_1) and (x_2, y_2) have the independent probability distributions $p_1(x_1, y_1)$ and $p_2(x_2, y_2)$, the probability distribution of the sums $z_1 = x_1 + y_1$ and $z_2 = x_2 + y_2$, may be expressed by the two-dimensional convolution defined by:

Figure 3.1 Illustration of a Monte Carlo simulation of M cables



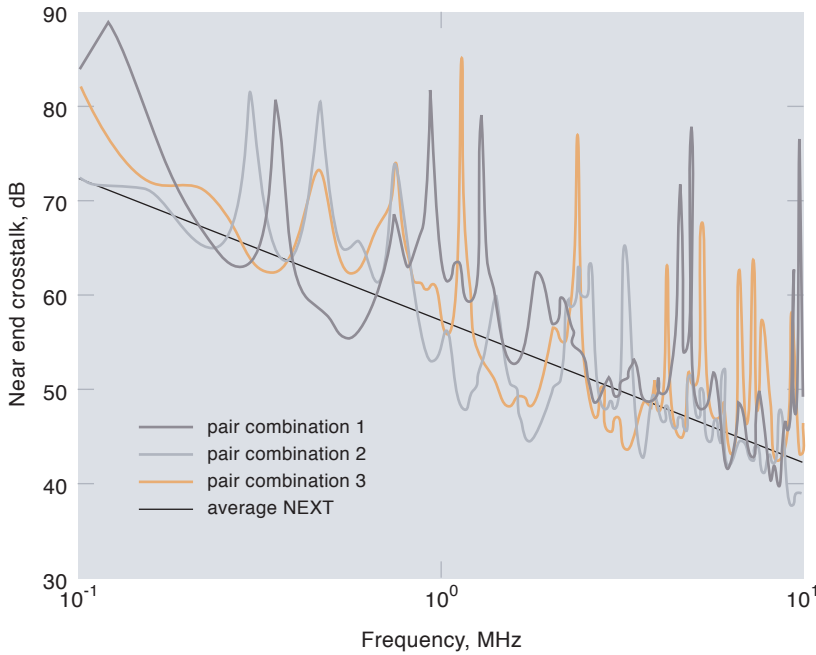


Figure 3.2 NEXT for three pair combinations generated by a Monte Carlo simulation

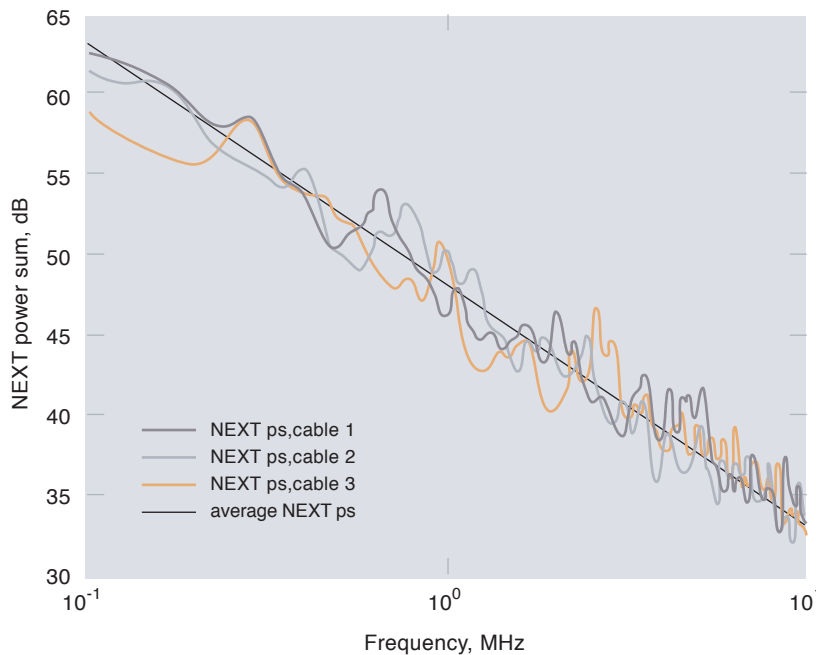


Figure 3.3 NEXT power sum for pair no. 1 of three different cables

$$\begin{aligned}
 p(z_1, z_2) &= p(x_1, x_2) * p(y_1, y_2) = \\
 &\int_{x_1=0}^{\infty} \int_{x_2=0}^{\infty} p(x_1, x_2) \cdot \\
 &p(z_1 - x_1, z_2 - x_2) dx_1 \cdot dx_2.
 \end{aligned} \tag{2.29}$$

3 Monte Carlo Simulation of NEXT

A cable is simulated in a computer by dividing the cable in short segments of length Δx , which is much shorter than the wavelength at the high-

est frequency under consideration ($\Delta x < \lambda / 20$). Statistically independent crosstalk couplings are drawn according to (2.2) for all cable segments and all pair combinations in each cable, and an ensemble of many different cables may be generated. This is illustrated in Figure 3.1.

Near end crosstalk for a specific pair combination is then calculated from Equation (2.1). This gives results as shown in Figure 3.2 for three individual pair combinations of the 10 pair cable (binder group) described in Table 2.1. This shows the typical peaky nature of NEXT that is observed in measurements. Average NEXT for this cable type is also shown in the figure as a broken line. In accordance with (2.7) this is a straight line with slope 15 dB/decade of the frequency.

The NEXT power sum in different cables of the type described in Table 2.1 has been calculated by simulation, and the result is shown in Figure 3.3 for pair no. 1 of three different cables. Due to the averaging over 9 different pair combinations, the NEXT power sum is less peaky than NEXT for individual pair combinations. The figure also shows the average NEXT power sum, which is a straight line that varies 15 dB/decade with frequency.

4 Channel Capacity

The theoretical channel capacity of a copper pair can be found by Shannon's channel capacity formula. A realistic estimate of the channel capacity of a pair may be calculated by inserting two factors into Shannon's channel capacity formula. This modified channel capacity will be a realistic estimate of the transmission rate of a multi-carrier system using adaptive modulation. According to Starr, Cioffi and Silverman [2], the transmission rate for a system operating in the frequency band $[f_l, f_h]$ of a cable may be expressed:

$$R(\ell) = k_{eff} \cdot \int_{f_l}^{f_h} \log_2 \left(1 + \eta \cdot \frac{\exp(-2\alpha\ell)}{z} \right) df. \tag{4.1}$$

$z = z(f)$ is the near end crosstalk power sum of the actual pair in the cable

$k_{eff} \leq 1$ is the ratio between user available bitrate and the total bitrate

$\eta = 10^{-\text{mdB}/10}$ where *mdB* is the distance to the Shannon bound in dB

The crosstalk power sum, z , is a random variable, and hence the channel capacity or transmission rate of the system will be a random variable. The moments and probability distribution of the channel capacity can either be calculated by analytical methods or by Monte Carlo simulation.

4.1 Calculation of Channel Capacity by Analytical Methods

The first and second order moments of the channel capacity can be calculated by means of the probability distributions of NEXT power sum found in Section 2.3. The average channel capacity will be:

$$E[R(\ell)] = k_{eff} \cdot \int_{f_l}^{f_h} \int_0^\infty p(z; f) \cdot \log_2 \left(1 + \eta \cdot \frac{\exp(-2\alpha\ell)}{z} \right) dz \cdot df. \quad (4.2)$$

The variance of the channel capacity will be:

$$\text{var}[R(\ell)] = E[R(\ell)^2] - \{E[R(\ell)]\}^2 \quad (4.3)$$

which may be expressed:

$$\begin{aligned} \text{var}[R(\ell)] = & k_{eff}^2 \cdot \int_{f_l}^{f_h} \int_{f_l}^{f_h} \int_0^\infty \int_0^\infty [p(z_1, z_2; f_1, f_2) - \\ & p(z_1; f_1) \cdot p(z_2; f_2)] \cdot \\ & \log_2 \left(1 + \eta \cdot \frac{\exp(-2\alpha_1\ell)}{z_1} \right) \\ & \log_2 \left(1 + \eta \cdot \frac{\exp(-2\alpha_2\ell)}{z_2} \right) dz_1 \cdot \\ & dz_2 \cdot df_1 \cdot df_2. \end{aligned} \quad (4.4)$$

The integrals of (4.2) and (4.4) cannot be solved analytically and have to be calculated either by approximations or numerical integration. In Section 5 it is shown that the channel capacity is approximately Gaussian, and hence an approximate analytical probability of the channel capacity is defined by (4.2) and (4.4). However, due to the complexity of both the above integrals and the convolution integral (2.28) the use of Monte Carlo simulation has been chosen in the rest of this paper.

4.2 Calculation of Channel Capacity by Monte Carlo Simulation

An ensemble of M different cables is generated in a Monte Carlo simulation. The channel capacity of cable no. k is found by:

$$R_k(\ell) = k_{eff} \cdot \int_{f_l}^{f_h} \log_2 \left(1 + \eta \cdot \frac{\exp(-2\alpha\ell)}{|H_{Nps}(f)|^2} \right) df, \quad (4.5)$$

where $|H_{Nps}(f)|^2$ is the NEXT power sum of this sum of this specific cable.

By means of the above formulas, the channel capacity can be calculated for the ensemble of cables. The average bitrate and the variance of the bitrate are estimated by:

$$R_{av}(\ell) = \frac{1}{M} \sum_{k=1}^M R_k(\ell), \quad (4.6)$$

$$\sigma_R^2(\ell) = \frac{1}{M} \sum_{k=1}^M R_k^2(\ell) - R_{av}^2(\ell). \quad (4.7)$$

The 1 % confidence limit (worst case estimate)

of the bitrate, $R_{1\%}(\ell)$, may be estimated from the ensemble $\{R_k(\ell)\}$.

5 System Example

Two-way transmission in pair cables is efficient only at the lowest frequencies. At higher frequencies it is clearly advantageous to use one-way transmission. This is exploited in the standards for ADSL [9] and VDSL [10]. The basic differences between one-way and two-way systems are explained for instance in a tutorial paper by Holte [11]. Two-way transmission is used in HDSL systems and in the newly standardised SHDSL system [1]. The advantages of calculating the channel capacity by this new approach are demonstrated by an example. This example is a system which uses a spectrum similar to that of the SHDSL system with maximum rate, 2.3 Mbit/s. The SHDSL system is a base-band system with a 10 dB bandwidth of approximately 400 kHz. The same bandwidth is used in the example, but because the crosstalk models are not valid below 100 kHz, the bandwidth has been shifted up to the frequency interval from 100 kHz to 500 kHz. Hence the capacity estimates will be pessimistic in comparison with a system operating from 0 – 400 kHz. The main assumptions of the system model used in the simulations are listed below:

Bandwidth:

100 kHz < f < 500 kHz

Modulation:

Multicarrier modulation (DMT), [12]

Modul. meth. in each sub-band:

Trellis coded M-QAM, $4 \leq M \leq 16384$

Corresp. bandwidth efficiency:

1 to 13 bit/s/Hz

Distance to Shannon bound:

9 dB (6 dB margin + 3 dB for TCM)

User available bitrate:

90 % of total bitrate

Cable type:

0.6 mm copper cable (22 AWG)

Cable size:

10 pair binder group

Attenuation constant:

15.1 dB/km at 1 MHz, proportional to \sqrt{f}

Phase constant:

31.4 rad/km at 1 MHz, proportional to f

Noise model:

Only NEXT between identical systems within a binder group; all pairs used

Average NEXT power sum:

47.9 dB at 1 MHz, proportional to $f^{1.5}$ (pair 1)

99 % conf.limit, NEXT ps:

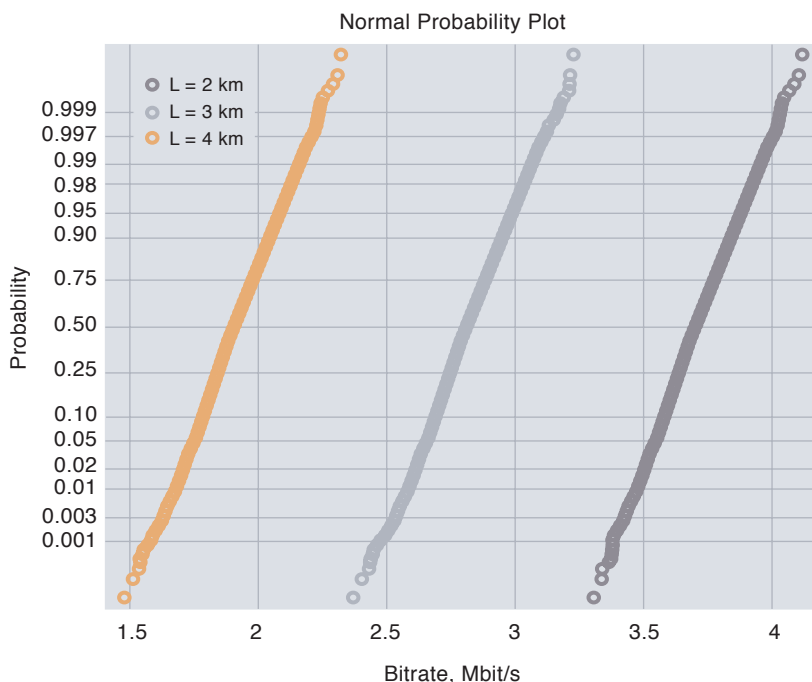
44.7 dB at 1 MHz (pair 1)

In order to take into account further practical limitations connected to adaptive modulation, the estimation of transmission rate has been slightly modified in comparison with Equation (4.5). The bandwidth efficiency is set to zero at frequencies where the signal to noise ratio gives a bandwidth efficiency below 1 bit/s/Hz, and the bandwidth efficiency is limited to a maximum of 13 bit/s/Hz. In the results presented in this paper, no quantisation of bandwidth efficiency has been used within the allowed interval. An alternative approach will be to quantise the bandwidth efficiency to integer values in order to simulate adaptive modulation which uses two-dimensional trellis coding. The reason for not using quantisation of bandwidth efficiency in this paper is that it introduces more or less random quantisation effects in the results, and this may be disturbing in relative comparisons. It is recommended to use quantisation of bandwidth efficiency in detailed dimensioning cases.

6 Results for Adaptive Modulation

Bitrates for different cable lengths have been estimated by means of the Monte Carlo simulation, based upon the above methods and the crosstalk measurements shown in Table 2.1. The estimated bitrates for 10,000 different cables at cable lengths 1, 2 and 3 km are shown in Figure 6.1 in a normal distribution plot. This result shows that the bitrate of a system with adaptive modulation fits very well with a Gaussian probability distribution.

Figure 6.1 Cumulative distribution of estimated bitrate for a 0.6 mm cable for different cable lengths



Both the average bitrate of the simulated cables and the 1 % confidence limit of the bitrate are shown in Figure 6.2 as a function of the cable length.

The difference between the two upper curves shows the average increase in bitrate that is obtained by using adaptive modulation instead of offering a guaranteed fixed rate. This increase is approximately 0.25 Mbit/s. The increase in bitrate is typically 10 % at range 3 km. Figure 6.2 also shows the bitrate for the same type of system based upon a traditional estimate of NEXT. This means that the 99 % confidence limit of NEXT power sum is used at all frequencies. The difference between the two lower curves shows the increase in bitrate that is obtained by using the capacity estimates of the new method instead of the traditional approach. The increase in bitrate is approximately 5 % for this example. The increase in bitrate between the two approaches decreases for longer ranges. The reason is that the useful bandwidth decreases with range due to high attenuation in the upper end of the frequency. Hence, the number of frequency bands that have independent NEXT will decrease as the cable length increases.

7 Implications for New Systems

The newly standardised SHDSL system [1] is a single carrier baseband system using 16-level pulse amplitude modulation and trellis coded modulation, and hence it has a bandwidth efficiency of less than 6 bit/s/Hz. If a new system is implemented which uses adaptive multicarrier modulation and is spectrally compatible with SHDSL, this new system will obtain significantly higher bitrates for the same cable lengths. The increase in bitrate is due to the following effects:

- higher bandwidth efficiency of multicarrier modulation, in particular at the lowest frequencies;
- less excess bandwidth for multicarrier modulation;
- independent NEXT in different frequency bands as explained in this paper;
- a variable rate system will have a larger average bitrate than a fixed rate system as shown in Figure 6.2.

8 Conclusions

It has been shown how the true channel capacity of a pair cable with respect to NEXT can be calculated both analytically and by simulation. It is demonstrated that by taking into account the frequency variations of near end crosstalk, the

channel capacity for two-way transmission in pair cables may be increased in comparison with traditional methods. The increase in worst-case transmission rate is approximately 5 % for a typical system example. For the same example it is also shown that the average bitrate for a rate adaptive system may be increased by typically 10 % in comparison with a system with a guaranteed minimum rate. This applies to multicarrier systems with adaptive modulation that use two-way transmission and echo cancelling at all frequencies.

Appendix A Probability Distributions of NEXT at a Single Frequency

The simplified notation $r = r_{i,j}(f)$ and $q = q_{i,j}(f)$ is used in this appendix for the real and imaginary part of the NEXT transfer function for one pair combination. Because the NEXT transfer function is a linear function of the coupling coefficients and the coupling coefficients are Gaussian with zero mean, r and q will be jointly Gaussian with zero mean. The variance of the real part will be given by:

$$E[r^2] = \beta_0^2 \int_0^\infty \int_0^\infty E[\kappa_{Ni,j}(x) \cdot \kappa_{Ni,j}(y)] \cdot \sin(2\beta x) \cdot \sin(2\beta y) \cdot \exp(-2\alpha x - 2\alpha y) dx \cdot dy. \quad (\text{A.1})$$

Inserting (2.2) gives:

$$E[r^2] = \beta_0^2 \cdot k_{Ni,j} \int_0^\infty \sin^2(2\beta x) \cdot \exp(-4\alpha x) dx. \quad (\text{A.2})$$

Solving the integral gives the result:

$$\sigma_r^2 = E[r^2] = \frac{\beta_0^2 \cdot k_{Ni,j}}{8\alpha} \cdot \frac{\beta^2}{\alpha^2 + \beta^2}. \quad (\text{A.3})$$

The remaining second order moments of the NEXT transfer function are calculated correspondingly:

$$E[r \cdot q] = \beta_0^2 \cdot k_{Ni,j} \int_0^\infty \sin(2\beta x) \cdot \cos(2\beta x) \cdot \exp(-4\alpha x) dx, \quad (\text{A.4})$$

$$E[r \cdot q] = \frac{\beta_0^2 \cdot k_{Ni,j}}{8} \cdot \frac{\beta}{\alpha^2 + \beta^2}, \quad (\text{A.5})$$

$$E[q^2] = \beta_0^2 \cdot k_{Ni,j} \int_0^\infty \cos^2(2\beta x) \cdot \exp(-4\alpha x) dx, \quad (\text{A.6})$$

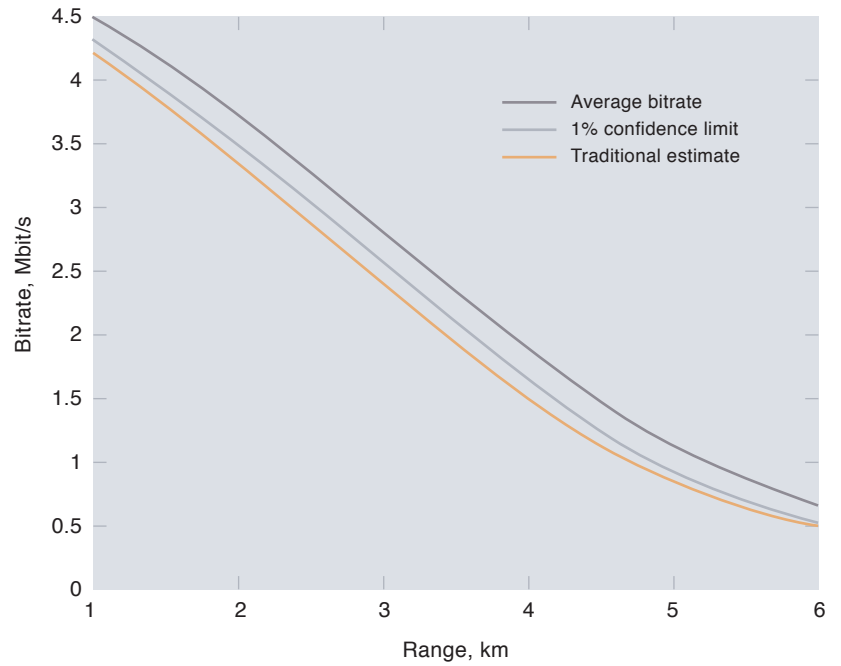


Figure 6.2 Estimated bitrates for two-way transmission as a function of range for a multicarrier system with adaptive modulation on a 0.6 mm pair cable

$$\sigma_q^2 = E[q^2] = \frac{\beta_0^2 \cdot k_{Ni,j}}{8\alpha} \cdot \frac{2\alpha^2 + \beta^2}{\alpha^2 + \beta^2}. \quad (\text{A.7})$$

The correlation coefficient is defined by:

$$\rho_{rq} = \frac{E[r \cdot q]}{\sigma_r \cdot \sigma_q} = \frac{\alpha}{\sqrt{2\alpha^2 + \beta^2}}. \quad (\text{A.8})$$

According to Papoulis [13], the probability density of NEXT for a single pair combination may be expressed:

$$p_{i,j}(r,q) = \frac{1}{2\pi \cdot \sigma_r \cdot \sigma_q \sqrt{1 - \rho_{rq}^2}} \cdot \exp\left[-\frac{1}{2(1 - \rho_{rq}^2)} \left(\frac{r^2}{\sigma_r^2} - 2\rho_{rq} \frac{r \cdot q}{\sigma_r \cdot \sigma_q} + \frac{q^2}{\sigma_q^2} \right)\right]. \quad (\text{A.9})$$

Insertion of (A.3), (A.7) and (A.8) and using the notation $P_{i,j} = P_{i,j}(f)$ for average NEXT transfer function gives:

$$p_{i,j}(r,q) = \frac{\sqrt{\alpha^2 + \beta^2}}{2\pi \cdot \beta \cdot P_{i,j}} \cdot \exp\left[-\frac{r^2 + q^2}{P_{i,j}} - \frac{2\alpha \cdot r}{\beta^2 \cdot P_{i,j}} (\alpha \cdot r - \beta \cdot q)\right]. \quad (\text{A.10})$$

By a change of variables the probability density may be expressed by the NEXT power transfer function $z = r^2 + q^2$ and the phase angle $\varphi = \arctan(q/r)$:

$$P_{i,j}(z, \varphi) = \frac{\sqrt{\alpha^2 + \beta^2}}{2\pi \cdot \beta \cdot P_{i,j}} \exp\left[-\frac{z}{P_{i,j}} \cdot \left(\frac{\alpha^2 + \beta^2}{\beta^2} + \frac{z \cdot \alpha}{\beta^2 \cdot P_{i,j}} \cdot \cos(2\varphi - \theta)\right)\right], \quad (\text{A.11})$$

where $\varphi = \arctan(\beta / \alpha)$, $z \geq 0$, and $0 \leq \varphi < 2\pi$.

The phase of the signals in different pairs in a cable is random. Hence, the phase angles of NEXT transfer functions are irrelevant. The probability density of the NEXT power transfer function for one pair combination is given by:

$$P_{i,j}(z) = \int_0^{2\pi} P(z, \varphi) d\varphi = \frac{\sqrt{\alpha^2 + \beta^2}}{\beta \cdot P_{i,j}} \exp\left[-\frac{z}{P_{i,j}} \cdot \frac{\alpha^2 + \beta^2}{\beta^2}\right] \cdot I_0\left(\frac{z \cdot \alpha \cdot \sqrt{\alpha^2 + \beta^2}}{\beta^2 \cdot P_{i,j}}\right), \quad z \geq 0 \quad (\text{A.12})$$

Under the assumption $\alpha \ll \beta$ this simplifies to an exponential distribution:

$$P_{i,j}(z) = \frac{1}{P_{i,j}} \exp\left(-\frac{z}{P_{i,j}}\right), \quad z \geq 0 \quad (\text{A.13})$$

Appendix B Joint Probability Distributions of NEXT at Two Different Frequencies

In order to calculate the joint probability distribution of NEXT at two different frequencies f_1 and f_2 , the following simplified notation is introduced in this appendix: $r_1 = r_{i,j}(f_1)$, $q_1 = q_{i,j}(f_1)$, $r_2 = r_{i,j}(f_2)$ and $q_2 = q_{i,j}(f_2)$. Furthermore, the parameters α , β , β_0 are given an additional index 1 or 2 to denote the values at f_1 and f_2 respectively. The vector of NEXT variables is defined by:

$$\mathbf{v}_{i,j} = [r_1 \ q_1 \ r_2 \ q_2]^T. \quad (\text{B.1})$$

where \top denotes transposition. The vector $\mathbf{v}_{i,j}$ will in correspondence with Appendix A be Gaussian with zero mean, and the probability density function is given by [13]:

$$P_{i,j}(\mathbf{v}_{i,j}; f_1, f_2) = \frac{1}{(2\pi)^2 \sqrt{|\mathbf{C}_{i,j}|}} \cdot \exp\left(-\frac{1}{2} \mathbf{v}_{i,j}^T \mathbf{C}_{i,j} \mathbf{v}_{i,j}\right), \quad (\text{B.2})$$

where $\mathbf{C}_{i,j}$ is the covariance matrix of $\mathbf{v}_{i,j}$. The elements of $\mathbf{C}_{i,j}$ are calculated in the same way as in Appendix A and are given by:

$$E[r_k \cdot r_l] = \beta_k \cdot \beta_l \cdot k_{Ni,j} \int_0^\infty \sin(2\beta_k x) \cdot \sin(2\beta_l x) \cdot \exp(-2\alpha_k x - 2\alpha_l x) dx, \quad (\text{B.3})$$

$$E[r_k \cdot q_l] = \beta_k \cdot \beta_l \cdot k_{Ni,j} \int_0^\infty \sin(2\beta_k x) \cdot \cos(2\beta_l x) \cdot \exp(-2\alpha_k x - 2\alpha_l x) dx, \quad (\text{B.4})$$

$$E[q_k \cdot q_l] = \beta_k \cdot \beta_l \cdot k_{Ni,j} \int_0^\infty \cos(2\beta_k x) \cdot \cos(2\beta_l x) \cdot \exp(-2\alpha_k x - 2\alpha_l x) dx, \quad k, l = 1, 2. \quad (\text{B.5})$$

The solutions are expressed:

$$E[r_k \cdot q_l] = \frac{\beta_{0k} \cdot \beta_{0l} \cdot k_{Ni,j}}{4} \left[\frac{\alpha_k + \alpha_l}{(\alpha_k + \alpha_l)^2 + (\beta_k - \beta_l)^2} + \frac{\alpha_k + \alpha_l}{(\alpha_k + \alpha_l)^2 + (\beta_k + \beta_l)^2} \right] \quad (\text{B.6})$$

$$E[r_k \cdot r_l] = \frac{\beta_{0k} \cdot \beta_{0l} \cdot k_{Ni,j}}{4} \left[\frac{\beta_k - \beta_l}{(\alpha_k + \alpha_l)^2 + (\beta_k - \beta_l)^2} + \frac{\beta_k + \beta_l}{(\alpha_k + \alpha_l)^2 + (\beta_k + \beta_l)^2} \right] \quad (\text{B.7})$$

$$E[q_k \cdot q_l] = \frac{\beta_{0k} \cdot \beta_{0l} \cdot k_{Ni,j}}{4} \left[\frac{\alpha_k + \alpha_l}{(\alpha_k + \alpha_l)^2 + (\beta_k - \beta_l)^2} + \frac{\alpha_k + \alpha_l}{(\alpha_k + \alpha_l)^2 + (\beta_k + \beta_l)^2} \right] \quad (\text{B.8})$$

Under the assumption $\alpha \ll \beta$ the last term in the equations (B.6) – (B.8) may be neglected, and for this case the covariance matrix is given by:

$$\mathbf{C}_{i,j} = \begin{bmatrix} a & 0 & c & d \\ 0 & a & -d & c \\ c & -d & b & 0 \\ d & c & 0 & b \end{bmatrix}. \quad (\text{B.9})$$

Using the notation $P_{1,i,j} = P_{i,j}(f_1)$ and $P_{2,i,j} = P_{i,j}(f_2)$ for average NEXT, the elements of the matrix may be expressed:

$$a = E[r_1^2] = E[q_1^2] = \frac{P_{1,i,j}}{2}, \quad (\text{B.10})$$

$$b = E[r_2^2] = E[q_2^2] = \frac{P_{2,i,j}}{2}, \quad (\text{B.11})$$

$$c = E[r_1 \cdot r_2] = E[q_1 \cdot q_2] = \frac{\sqrt{P_{1,i,j} \cdot P_{2,i,j}} \cdot (\alpha_1 + \alpha_2) \cdot \sqrt{\alpha_1 \cdot \alpha_2}}{2 \cdot (\alpha_1 + \alpha_2)^2 + (\beta_1 - \beta_2)^2}, \quad (\text{B.12})$$

$$d = E[r_1 \cdot q_2] = -E[q_1 \cdot r_2] = \frac{\sqrt{P_{1,i,j} \cdot P_{2,i,j}} \cdot (\beta_1 - \beta_2) \cdot \sqrt{\alpha_1 \cdot \alpha_2}}{2 \cdot (\alpha_1 + \alpha_2)^2 + (\beta_1 - \beta_2)^2}, \quad (\text{B.13})$$

$$E[r_1 \cdot q_1] = E[r_2 \cdot q_2] = 0. \quad (\text{B.14})$$

Gibbs and Addie [3] have neglected the term d . This is obviously a significant term, which means that Gibbs and Addie's calculation of joint probability densities at two frequencies is incorrect.

The determinant of $C_{i,j}$ will be:

$$|C_{i,j}| = (a \cdot b - c^2 - d^2)^2 = g^2. \quad (\text{B.15})$$

where g may be expressed:

$$g = a \cdot b - c^2 - d^2 = \frac{P_{1,i,j} \cdot P_{2,i,j}}{4} \cdot \left(1 - \frac{\alpha_1 \cdot \alpha_2}{(\alpha_1 + \alpha_2)^2 + (\beta_1 - \beta_2)^2} \right). \quad (\text{B.16})$$

The inverse correlation matrix is given by:

$$C_{i,j}^{-1} = \frac{1}{g} \begin{bmatrix} b & 0 & -c & -d \\ 0 & b & d & -c \\ -c & d & a & 0 \\ -d & -c & 0 & a \end{bmatrix}. \quad (\text{B.17})$$

The probability density of $v_{i,j}$ may hence be expressed:

$$P_{i,j}(v_{i,j}) = \frac{1}{(2\pi)^2 g} \cdot \exp\left(-\frac{1}{2g} \left[b(r_1^2 + q_1^2) + a(r_2^2 + q_2^2) - 2c(r_1 \cdot r_2 + q_1 \cdot q_2) - 2d(r_1 \cdot q_2 - q_1 \cdot r_2) \right] \right). \quad (\text{B.18})$$

A change of variables to NEXT power transfer functions and NEXT phase angles is carried out. The new variables are, $z_1 = r_1^2 + q_1^2$, $\varphi = \arctan[q_1 / r_1]$, $z_2 = r_2^2 + q_2^2$, and $\theta = \arctan[q_2 / r_2]$, and the probability density may then be written:

$$P_{i,j}(z_1, \varphi, z_2, \theta) = \frac{1}{(4\pi)^2 g} \cdot \exp\left(-\frac{1}{2g} \left[bz_1 + az_2 + 2\sqrt{z_1 \cdot z_2} \cdot \sqrt{c^2 + d^2} \cos(\varphi - \theta + \psi) \right] \right), \quad (\text{B.19})$$

where $\psi = \arctan[d / c]$, $z_1, z_2 \geq 0$, and $0 \leq \varphi, \theta < 2\pi$. The phase angles of NEXT transfer functions are irrelevant, and the joint probability density of z_1 and z_2 will be:

$$P_{i,j}(z_1, z_2) = \frac{1}{2\pi \cdot 2\pi} \int_0^{2\pi} \int_0^{2\pi} P_{i,j}(z_1, \varphi, z_2, \theta) d\varphi \cdot d\theta = \frac{1}{4g} \cdot \exp\left(-\frac{bz_1 + az_2}{2g}\right) \cdot I_0\left(\frac{\sqrt{z_1 \cdot z_2} \cdot \sqrt{c^2 + d^2}}{g}\right), \quad (\text{B.20})$$

$$z_1, z_2 \geq 0.$$

The parameters a, b, c, d are replaced according to (B.10) – (B.13), and the result is expressed by the normalised covariance ρ defined in (2.14).

Hence, the joint probability density of the NEXT power transfer function at two frequencies f_1 and f_2 for one pair combination is given by:

$$P_{i,j}(z_1, z_2) = \frac{1}{P_{1,i,j} \cdot P_{2,i,j} \cdot (1 - \rho)} \cdot \exp\left(-\frac{z_1}{P_{1,i,j}} - \frac{z_2}{P_{2,i,j}}\right) \cdot I_0\left(\frac{2\sqrt{\rho}}{1 - \rho} \sqrt{\frac{z_1 \cdot z_2}{P_{1,i,j} \cdot P_{2,i,j}}}\right), \quad (\text{B.21})$$

$$z_1, z_2 \geq 0.$$

References

- 1 ITU-T. *Single-Pair High-Speed Digital Subscriber Line (SHDSL) transceivers*. Geneva, 2001. (ITU-T Recommendation G.991.2.)
- 2 Starr, T, Cioffi, J M, Silverman, P J. *Understanding digital subscriber line technology*. Upper Saddle River, Prentice Hall, 1999.
- 3 Gibbs, A J, Addie, R. The covariance of near end crosstalk and its application to PCM system engineering in multipair cable. *IEEE Trans. on Commun.*, COM-27 (2), 1979, 469–477.
- 4 Klein, W. *Die Theorie des Nebensprechens auf Leitungen*. Berlin, Springer, 1955.
- 5 Cravis, H, Crater, T V. Engineering of T1 carrier system repeatered lines. *Bell System Techn. Journal*, March 1963, 431–486.
- 6 Bradley, S D. Crosstalk considerations for a 48 channel PCM repeatered line. *IEEE Trans. on Communications*, COM-23 (7), 1975, 722–728.
- 7 Holte, N. A Crosstalk Model for Cross-Stranded Cables. *Int. Wire & Cable Symp.*, Cherry Hill, USA, Nov. 1982.
- 8 Holte, N. *Crosstalk in subscriber cables, Final report*. Trondheim, SINTEF, 1985.

- (SINTEF report no. STF44 F85001.) (In Norwegian.)
- 9 ITU-T. *Asymmetric Digital Subscriber Line (ADSL) transceivers*. Geneva, 1999. (ITU-T Recommendation G.992.1.)
 - 10 ETSI. *Very high speed Digital Subscriber Line (VDSL); Part 1: Functional requirements*. Sophia Antipolis, 1999. (ETSI TS 101 270-1, V1.2.1.)
 - 11 Holte, N. *Broadband communication in existing copper cables by means of xDSL systems – a tutorial*. NORSIG, Trondheim, Norway, October 2001.
 - 12 Bingham, J A C. Multicarrier modulation for data transmission : An idea whose time has come. *IEEE Communications Magazine*, May 1990, 5–19.
 - 13 Papoulis, A. *Probability, random variables, and stochastic processes*. 3rd. ed., New York, McGraw Hill, 1991.
 - 14 Abramowitz, M, Stegun, I A. *Handbook of mathematical functions*. New York, Dover, 1970.

Bounds on the Average Spectral Efficiency of Adaptive Coded Modulation

KJELL J. HOLE



Kjell Jørgen Hole (41) received his BSc, MSc and PhD degrees in computer science from the University of Bergen in 1984, 1987 and 1991, respectively. He is currently Senior Research Scientist at the Department of Telecommunications at the Norwegian University of Science and Technology (NTNU) in Trondheim. His research interests are in the areas of coding theory and wireless communications.

Kjell.Hole@ii.uib.no

An introduction to adaptive coded modulation (ACM) was given in an earlier paper by Hole and Øien [4]. It was shown that ACM may achieve a large average spectral efficiency (ASE) on wireless channels with slowly varying frequency-flat fading. This paper presents new upper bounds on the ASE of ACM with maximum likelihood decoding and sequential decoding. The bounds are compared to the theoretical maximum ASE. It is found that this theoretical maximum provides an optimistic upper bound on the achievable ASE of practical ACM schemes. However, the new bounds indicate that ACM may still provide a large ASE.

I. Introduction

Time-varying channel conditions is an important feature of most wireless communication systems. Future systems must therefore exhibit a high degree of *adaptivity* on many levels to support traffic flows with large information rates (measured in transmitted information bits per second). Examples of such adaptivity are: power control, code adaptation, bandwidth adaptation, antenna adaptation, and protocol adaptation [1], [2]. This paper deals with code adaptation.

The *spectral efficiency* of a wireless channel or link is equal to the information rate per unit bandwidth. When the instantaneous spectral efficiency varies due to code adaptation, the *average spectral efficiency* (ASE) should be maximized. Goldsmith and Varaiya [3] have determined the *maximum ASE* of a single-user communication system with a slowly varying frequency-flat fading channel, fixed transmit signal power, and perfect *channel-state information* (CSI) available at the transmitter and receiver.

It is shown in [3] that the maximum ASE may be obtained by a theoretical *adaptive coding* scheme. *Adaptive coded modulation* (ACM) with fixed transmit power may be a practical adaptive transmission technique for frequency-flat fading channels with available CSI [4]–[11]. ACM may utilize e.g. a set of classical trellis codes [12]–[17] or a set of turbo-trellis codes [18]–[21] in which each code is designed for good performance on an additive white Gaussian noise (AWGN) channel. If the codes in a set are based on quadrature amplitude modulation (QAM) constellations of different sizes [22], then a low bit error rate (BER) and large ASE may be obtained simultaneously by switching adaptively between the codes according to the CSI. The goal of this paper is to present new upper bounds on the ASE of single-user wireless ACM/QAM channels with frequency-flat fading.

We begin by introducing, in Section II, a communication system model utilizing ACM/QAM on an arbitrary single-user flat-fading channel. Section III then upper bounds the ASE of ACM/QAM both for maximum likelihood decoding (MLD) and for sequential decoding (SD). As an example, the MLD bound is applied to single-user Nakagami fading channels, and both the MLD and SD bounds are compared to the theoretical maximum ASE for Nakagami fading channels [23]. Conclusions are drawn in Section IV.

II. Fading Channels and ACM

The discrete-time system model consists of a transmitter and a receiver communicating over a single user wireless channel degraded by multipath fading. The fading is assumed to remain nearly constant over hundreds of channel symbols. Pilot symbols are sent repeatedly over the channel to ensure that the receiver is able to fully compensate for the amplitude and phase variations in the received signal, i.e. we assume ideal channel estimation and coherent detection.¹⁾ Hence, we may model the stationary and ergodic fading amplitude $\alpha(t) (\geq 0)$ as a stochastic variable with *real* values.

Let the transmitted signal have complex baseband representation $x(t)$ at time index $t \in \{0, 1, 2, \dots\}$. The received baseband signal is then given by $y(t) = \alpha(t)x(t) + n(t)$ where $n(t)$ denotes complex AWGN. The real and imaginary parts of the noise are statistically independent, both with variance $(N_0B)/2$ where N_0 [W/Hz] is the total one-sided noise power spectral density and B [Hz] is the one-sided channel bandwidth.

Denote the average transmit power by S [W]. The instantaneous received *carrier-to-noise ratio* (CNR) is represented by the stochastic variable $\gamma(t) = \alpha^2(t)S/(N_0B)$. (In the sequel we omit the time reference t and refer to α and γ .) Let $p(\gamma)$ be the probability density function (pdf)

¹⁾ The overhead bandwidth associated with the pilot symbols is ignored.

of the CNR γ . We only assume that $p(\gamma)$ is a *continuous function* on the interval $[0, \infty)$. The CNR has expectation $E[\gamma] = \bar{\gamma} = \Omega S / (N_0 B)$, where $E[\alpha^2] = \Omega$ is the average received power gain. We assume that there exists a noiseless and zero-delay feedback channel from the receiver to the transmitter such that both the transmitter and receiver have perfect knowledge of the instantaneous received CNR γ at all times.

Let N quantization levels (or *fading regions*) represent the time-varying received CNR γ . When ACM is used, one trellis code designed to combat AWGN is assigned to each fading region [4]–[11]. The N fading regions are defined by the thresholds $0 < \gamma_1 < \gamma_2 < \dots < \gamma_{N+1} = \infty$. Code n , $n \in \{1, 2, \dots, N\}$, is utilized every time the instantaneous received CNR γ falls in region n , i.e. when $\gamma_n \leq \gamma < \gamma_{n+1}$. The ACM system is designed such that the BER never exceeds a given target maximum BER₀ for any CNR γ . Details on how to choose thresholds $\{\gamma_n\}_{n=1}^{N+1}$ to achieve BER \leq BER₀ for a given set of N codes can be found in [9].

Fading region $n = 1$ represents the smallest values of γ for which information is transmitted. No information is sent when $\gamma < \gamma_1$ simply because the channel quality is too bad to successfully transmit any information with the available codes. Hence, an ACM scheme experiences an *outage* during which information must be buffered at the transmitter end. The probability of outage is

$$P^{out}(\gamma_1) = \int_0^{\gamma_1} p(\gamma) d\gamma = 1 - \int_{\gamma_1}^{+\infty} p(\gamma) d\gamma. \quad (1)$$

III. Bounds on ASE of ACM

This section upper bounds the ASE of ACM utilizing trellis codes for QAM signal constellations.

III.A Bound for MLD

Let $4 \leq M_1 < M_2 < \dots < M_N$ denote the number of symbols in N two-dimensional QAM constellations of growing size, and let code n be based on the constellation with M_n symbols. For some small fixed $L \in \{1, 2, \dots\}$, the encoder for code n accepts $L \cdot \log_2(M_n) - c$ information bits at each time index $k = L \cdot t \in \{0, L, 2L, \dots\}$ and generates $L \cdot \log_2(M_n)$ coded bits, $1 \leq c < L \cdot \log_2(M_n)$. The coded bits specify L modulation symbols in the n th QAM constellation. These symbols are transmitted at time indices $k, k + 1, \dots, k + L - 1$. The L two-dimensional symbols can be viewed as one $2L$ -dimensional symbol, and for this reason the code is said to be a $2L$ -dimensional trellis code.

Assuming Nyquist signaling, the time used to transmit one two-dimensional QAM symbol is $T_s = 1/B$ [s]. Since the number of information bits per QAM symbol is $\log_2(M_n) - c/L$, the information rate of code n is $R_n = (\log_2(M_n) - c/L)/T_s$ [bits/s], and the spectral efficiency is $R_n/B = \log_2(M_n) - c/L$ [bits/s/Hz]. Consequently, the ASE of all N codes is [9],[24]

$$\begin{aligned} \text{ASE}(\{\gamma_n\}) &= \sum_{n=1}^N \frac{R_n}{B} \cdot P(\gamma_n, \gamma_{n+1}) \\ &= \sum_{n=1}^N (\log_2(M_n) - c/L) P(\gamma_n, \gamma_{n+1}) \quad (2) \\ &\text{[bits/s/Hz]} \end{aligned}$$

where

$$P(\gamma_n, \gamma_{n+1}) = \int_{\gamma_n}^{\gamma_{n+1}} p(\gamma) d\gamma \quad (3)$$

is the probability of the instantaneous CNR γ falling in fading region n .

In principle, infinitely many sets of $2L$ -dimensional codes are available. We restrict ourselves to code sets which contain N codes and consider all such sets having the same parameters c, L , and $\{M_n\}_{n=1}^N$. In general, the code sets will correspond to different sets of thresholds $\{\gamma_n\}$ for a common target BER₀ [9]. A change of code set therefore implies changing one or more of the associated thresholds. To determine which code set maximizes the ASE under the given conditions, we first prove the following result.

Lemma 1: For given c, L, BER_0 , and $\{M_n\}$, the ASE defined by (2) and (3) does not decrease but may increase when a threshold γ_n is reduced.

Proof: First, let $n > 1$ and reduce the n th threshold from γ_n to $\hat{\gamma}_n$ such that $\gamma_{n-1} \leq \hat{\gamma}_n < \gamma_n$.

Keep the other thresholds fixed. We study the sum of the $(n-1)$ th and n th terms in (2). Before the threshold is changed, the sum is given by

$$Q = \tilde{R}_{n-1} \cdot P(\gamma_{n-1}, \gamma_n) + \tilde{R}_n \cdot P(\gamma_n, \gamma_{n+1})$$

for $\tilde{R}_j = R_j / B = \log_2(M_j) - c/L, j = n-1, n$,

while after the threshold is changed the sum becomes

$$\hat{Q} = \tilde{R}_{n-1} \cdot P(\gamma_{n-1}, \hat{\gamma}_n) + \tilde{R}_n \cdot P(\hat{\gamma}_n, \gamma_{n+1}).$$

The pdf $p(\gamma)$ in (3) is assumed to be a continuous function and we may write

$$\begin{aligned} P(\gamma_{n-1}, \gamma_n) &= P(\gamma_{n-1}, \hat{\gamma}_n) + P(\hat{\gamma}_n, \gamma_n) \\ P(\hat{\gamma}_n, \gamma_{n+1}) &= P(\hat{\gamma}_n, \gamma_n) + P(\gamma_n, \gamma_{n+1}) \end{aligned}$$

The difference $D = \hat{Q} - Q$ then becomes

$$D = [\tilde{R}_n - \tilde{R}_{n-1}] \cdot P(\hat{\gamma}_n, \gamma_n)$$

where $\tilde{R}_n > \tilde{R}_{n-1} > 0$. Since $P(\hat{\gamma}_n, \gamma_n) \geq 0$, we have $D \geq 0$. When $n = 1$ the difference reduces to

$$\begin{aligned} D &= \tilde{R}_1 [P(\hat{\gamma}_1, \gamma_2) - P(\gamma_1, \gamma_2)] \\ &= \tilde{R}_1 \cdot P(\hat{\gamma}_1, \gamma_1) \end{aligned}$$

for some $\tilde{R}_1 > 0$. Since $P(\hat{\gamma}_1, \gamma_1) \geq 0$, we again have $D \geq 0$. Q.E.D.

It is possible to lower bound the thresholds $\{\gamma_n\}$ which can be used for a given BER_0 . Code n is only active when the instantaneous CNR falls in fading region n , i.e. $\gamma \in [\gamma_n, \gamma_{n+1})$. Hence, for a given region n , the fading channel can be approximated by a bandlimited AWGN channel with CNR at least equal to γ_n . The BER must never exceed the target maximum BER_0 and code n must therefore achieve $\text{BER} \leq \text{BER}_0$ on an AWGN channel of CNR γ_n [9]. The maximum spectral efficiency of this AWGN channel is equal to the Shannon capacity, C_n [bits/s], divided by the bandwidth B [Hz]

$$\frac{C_n}{B} = \log_2(1 + \gamma_n) \text{ [bits/s/Hz]}.$$

Since the spectral efficiency of code n must satisfy $R_n/B \leq C_n/B$, we have

$$\gamma_n \geq \frac{M_n}{2^{c/L}} - 1 \stackrel{\text{def}}{=} \gamma_n^{\text{MLD}}, \quad n = 1, 2, \dots, N.$$

Note that the spectral efficiency R_n/B is obtained for a small finite target BER_0 while an arbitrarily small BER is assumed for the maximum spectral efficiency C_n/B . Hence, the lower bound γ_n^{MLD} is in fact valid for any fixed BER_0 . Furthermore, observe that the proof of the Shannon capacity C_n (see e.g. [25, Sec. 10.1]) does not require MLD; however, we use the superscript MLD to denote that γ_n^{MLD} is the minimum possible threshold also for MLD.

As an example, let the number of symbols in the n th QAM constellation be $M_n = 2^{n+1}$. The spectral efficiency $R_n/B = n + 1 - c/L$ and the minimum threshold $\gamma_n^{\text{MLD}} = 2^{n+1-c/L} - 1$ are listed (in dB) in Table 1 for $c = 1$, $L = 1, 2$, and $n = 1, 2, \dots, 11$.

Theorem 1: For given c , L , and $\{M_n\}$, the ASE defined by (2) and (3) is maximized for the minimum thresholds γ_n^{MLD} , $n = 1, 2, \dots, N$.

n	M_n	L	R_n/B	γ_n^{MLD} [dB]	γ_n^{SD} [dB]
1	4	1	1.0	0.00	2.39
		2	1.5	2.62	4.80
2	8	1	2.0	4.77	6.80
		2	2.5	6.68	8.61
3	16	1	3.0	8.45	10.30
		2	3.5	10.13	11.94
4	32	1	4.0	11.76	13.53
		2	4.5	13.35	15.09
5	64	1	5.0	14.91	16.63
		2	5.5	16.46	18.17
6	128	1	6.0	17.99	19.69
		2	6.5	19.52	21.21
7	256	1	7.0	21.04	22.73
		2	7.5	22.55	24.24
8	512	1	8.0	24.07	25.75
		2	8.5	25.58	27.26
9	1024	1	9.0	27.08	28.76
		2	9.5	28.59	30.27
10	2048	1	10.0	30.10	31.78
		2	10.5	31.61	33.28
11	4096	1	11.0	33.11	34.79
		2	11.5	34.62	36.30

Proof: The ASE defined by (2) and (3) has an absolute maximum at $(\gamma_1^{\text{MLD}}, \dots, \gamma_n^{\text{MLD}})$. To see this, reduce the threshold γ_1 to its minimum value γ_1^{MLD} . It follows from Lemma 1 that the ASE is not reduced. Repeat the procedure for each of the remaining thresholds $\gamma_2, \gamma_3, \dots, \gamma_N$. Q.E.D.

Since the outage probability $P^{\text{out}}(\gamma_1)$ in (1) does not increase but may decrease when the threshold γ_1 is reduced, we also have

Corollary 1: The minimum outage probability in an ACM/QAM system is $P^{\text{out}}(\gamma_1^{\text{MLD}})$.

To obtain an adaptive codec with ASE close to the maximum given by Theorem 1, the information rate R_n of each trellis code n must be close to the Shannon capacity C_n of an AWGN channel with CNR equal to γ_n^{MLD} . The capacity C_n is obtained with a Gaussian distribution over continuous-valued channel input symbols. ACM utilizes trellis codes based on equiprobable discrete-valued QAM symbols. However, there exist various techniques, called constellation shaping techniques, that achieve a more Gaussian-like distribution of the coded QAM symbols [17].

Table 1 Spectral efficiency R_n/B [bits/s/Hz] and minimum thresholds γ_n^{MLD} [dB] and γ_n^{SD} [dB] for $c = 1$ and various values of n and L

ASE bound for MLD, $c=L=1$

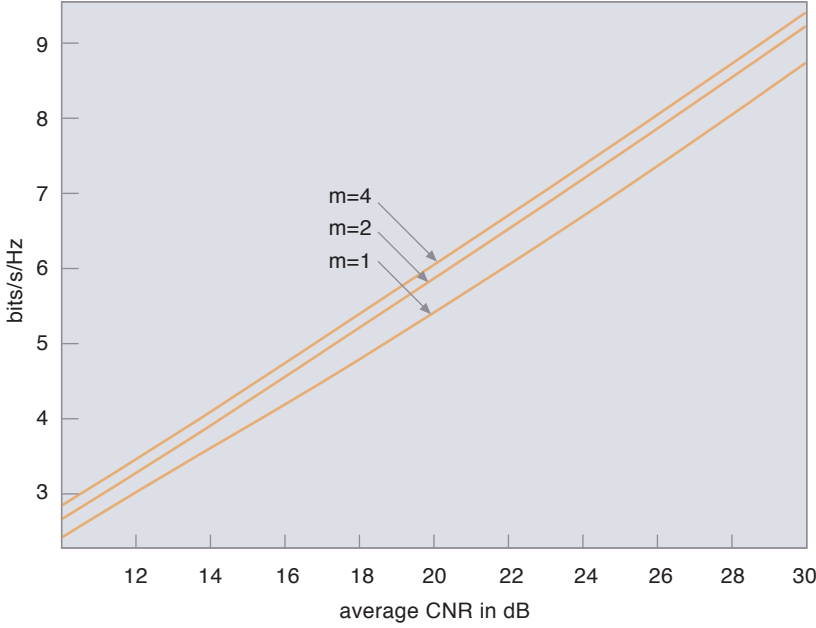
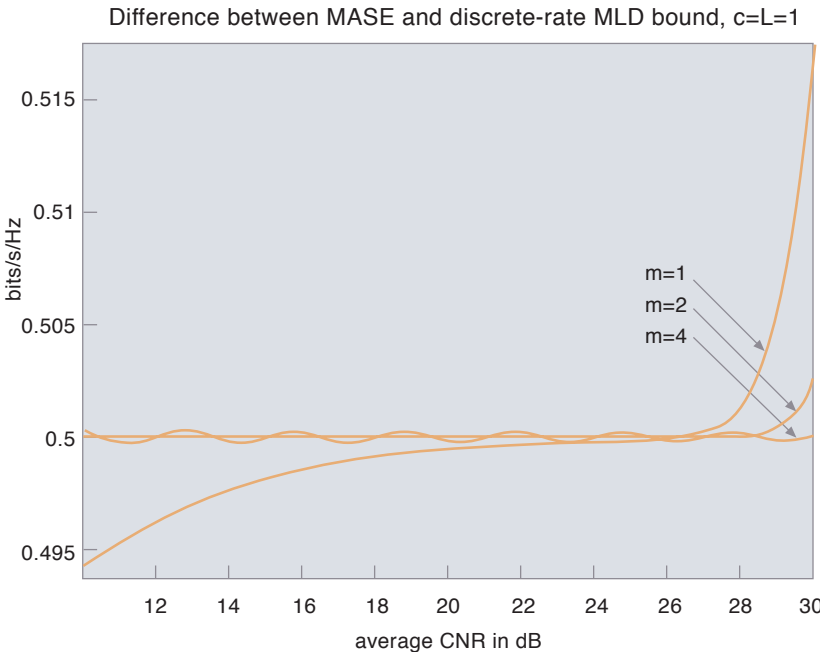


Figure 1 Maximum achievable ASE of ACM/QAM obtained from (2) and (4) for minimum thresholds γ_n^{MLD} in Table 1, $c = L = 1$, $n = 1, 2, \dots, 11$

As an example, we apply Theorem 1 to maximize the ASE of Nakagami multipath fading channels whose fading amplitude α has a Nakagami pdf controlled by a real parameter $m = \Omega^2/\text{Var}[\alpha^2]$ with constraint $m \geq 1/2$ [26, p. 48].²⁾ Here, $\text{Var}[\cdot]$ denotes the variance. The probability of γ falling in the n th fading region (3) is equal to [24]

$$P(\gamma_n, \gamma_{n+1}) = \frac{\Gamma\left(m, \frac{m\gamma_n}{\gamma}\right) - \Gamma\left(m, \frac{m\gamma_{n+1}}{\gamma}\right)}{\Gamma(m)}, \quad (4)$$

Figure 2 Difference between theoretical maximum ASE and maximum achievable ASE of ACM/QAM



where $\Gamma(\cdot, \cdot)$ is the complementary incomplete gamma function [27, Eq. (8.350.2)] and $\Gamma(m)$ is the gamma function, which equals $\Gamma(m) = (m - 1)!$ when m is a positive integer.

Let $m \in \{1, 2, 4\}$, $c = L = 1$, and $M_n = 2^{n+1}$ for $n = 1, \dots, 11$. From Theorem 1, the maximum achievable ASE of ACM/QAM is obtained from (2) and (4) using the $N = 11$ minimum thresholds γ_n^{MLD} for $L = 1$ in Table 1. The maximum achievable ASE is plotted in Figure 1. Observe that a large ASE is achievable even for Rayleigh fading ($m = 1$). The difference between the theoretical maximum ASE [23, Eq. (23)], often denoted MASE, and the maximum achievable ASE of ACM/QAM is plotted in Figure 2. The curves show that the use of $N = 11$ codes (with Shannon capacity performance on continuous input AWGN channels) results in a maximum achievable ASE of about .5 bit/s/Hz less than the theoretical maximum ASE. Nearly the same difference was found for $L = 2$.

III.B Bound for Sequential Decoding

In practice, MLD with the Viterbi algorithm is only implemented for classical QAM trellis codes with relatively short constraint lengths. For a set of codes with large constraint lengths, we may use SD instead [28]–[30]. The channel cutoff rate for each code n , $R_0(\gamma_n)$ [bits/s/Hz], is then the maximum spectral efficiency at which the average number of computations per decoded information bit is bounded for the lower threshold γ_n . For the bandlimited complex AWGN channel with Gaussian distributed input, we have [28]

$$R_0(\gamma_n) = (\log_2 e) \left[1 + \frac{\gamma_n}{2} - \sqrt{1 + \left(\frac{\gamma_n}{2}\right)^2} \right] + \log_2 e \left[\frac{1}{2} \left(1 + \sqrt{1 + \left(\frac{\gamma_n}{2}\right)^2} \right) \right] \quad [\text{bits/s/Hz}].$$

Theorem 2: If SD with a finite average number of computations per decoded bit is to be used, then the ASE defined by (2) and (3) is maximized for the thresholds

$$\gamma_n^{\text{SD def}} = \min \left\{ \gamma_n \mid R_0(\gamma_n) \geq \log_2(M_n) - c/L \right\} \quad (5)$$

$n = 1, 2, \dots, N,$

given c , L , and $\{M_n\}$.

Proof: The theorem is an immediate consequence of Lemma 1 and the following observation: When a set of N trellis codes based on QAM constellations with M_n symbols are used

²⁾ Nakagami is a general fading distribution that reduces to Rayleigh for $m = 1$. It also approximates the Rician distribution for $m > 1$, and it can approach the log-normal distribution.

in conjunction with SD, the spectral efficiency of code n given by $R_n/B = \log_2(M_n) - c/L$ must satisfy $R_n/B \leq R_0(\gamma_n)$ where γ_n is the lower threshold of fading region n . Q.E.D.

Corollary 2: The minimum outage probability for SD is $P_{out}(\gamma_1^{SD})$.

A variation on Newton's method was used to determine the thresholds γ_n^{SD} in (5). These values are tabulated in the rightmost column of Table 1 for $c = 1$, $L = 1, 2$ and $M_n = 2^{n+1}$, $n = 1, \dots, 11$. For two-dimensional codes ($L = 1$) and Nakagami fading with $m \in \{1, 2, 4\}$, the difference between the theoretical maximum ASE [23, Eq. (23)] and achievable ASE (2), (4) is plotted in Figure 3. Nearly 1.06 bits/s/Hz is lost compared to the theoretical maximum ASE. The same is true for four-dimensional codes ($L = 2$). Consequently, assuming optimal performing codes, an increase in ASE of about .56 bit/s/Hz may be obtained by utilizing MLD instead of SD where MLD is feasible.

IV. Conclusions

For single-user systems with frequency-flat fading, it seems clear that the theoretical maximum ASE [3], [23] provides an optimistic upper bound on the achievable ASE of practical ACM codecs. It was found that any sets of two-dimensional or four-dimensional trellis codes for MLD have ASE at least .5 bit/s/Hz less than the maximum ASE (see Figure 2) on a Nakagami fading channel. However, ACM may still provide a large ASE.

The ASE of optimal performing ACM with MLD is about .56 bit/s/Hz larger than the ASE of optimal performing ACM with SD. A preliminary investigation indicates that this difference may be smaller for sets of known trellis codes. Hence, ACM with SD may be an interesting alternative to ACM with MLD. More research is needed to determine the real world performance of ACM with SD on wireless channels.

References

- 1 Meyr, H. Algorithm design and system implementation for advanced wireless communications systems. In: *Proc. International Zurich Seminar on Broadband Communications (IZS'2000)*, Zürich, Switzerland, Feb. 2000.
- 2 Bose, V, Wetherall, D, Gutttag, J. Next century challenges: RadioActive Networks. In: *Proc. ACM/IEEE International Conference on Mobile Computing and Networking (MOBICOM'99)*, Seattle, WA, Aug. 1999.
- 3 Goldsmith, A J, Varaiya, P P. Capacity of fading channels with channel side informa-

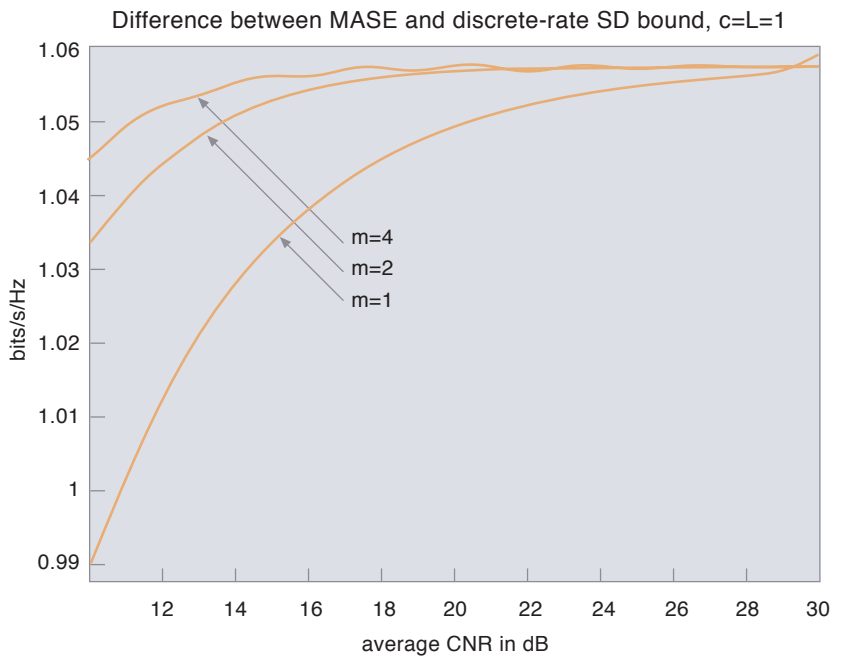


Figure 3 Difference between theoretical maximum ASE and achievable ASE of ACM/QAM with SD, obtained from (2) and (4) for minimum thresholds γ_n^{SD} in Table 1, $c = L = 1$, $n = 1, 2, \dots, 11$

- tion. *IEEE Trans. Inform. Theory*, 43 (6), 1986–1992, 1997.
- 4 Hole, K J, Øien, G E. Adaptive coding and modulation: A key to bandwidth-efficient multimedia communications in future wireless systems. *Teletronikk*, 97 (1), 49–57, 2001.
- 5 Goldsmith, A J, Chua, S-G. Adaptive coded modulation for fading channels. *IEEE Trans. Commun.*, 46 (5), 595–602, 1998.
- 6 Lau, V K N, Macleod, M D. Variable rate adaptive trellis coded QAM for high bandwidth efficiency applications in Rayleigh fading channels. In: *Proc. 48th IEEE Vehicular Technology Conference (VTC'98)*, Ottawa, Canada, May 1998, 348–352.
- 7 Goldsmith, A J. Adaptive modulation and coding for fading channels. In: *Proc. IEEE Inform. Theory and Commun. Workshop*, Kruger National Park, South Africa, June 1999, 24–26.
- 8 Goeckel, D L. Adaptive coding for time-varying channels using outdated fading estimates. *IEEE Trans. Commun.*, 47 (6), 844–855, 1999.
- 9 Hole, K J, Holm, H, Øien, G E. Adaptive multidimensional coded modulation over flat fading channels. *IEEE J. Select. Areas Commun.*, 18 (7), 1153–1158, 2000.

- 10 Hole, K J, Øien, G E. Spectral efficiency of adaptive coded modulation in urban micro-cellular networks. *IEEE Trans. Veh. Technol.*, 50 (1), 205–222, 2001.
- 11 Vishwanath, S, Goldsmith, A J. Exploring adaptive turbo coded modulation for flat fading channels. In: *Proc. 52nd IEEE Vehicular Technology Conference (2000 VTC-Fall)*, Boston, MA, Sept. 2000.
- 12 Ungerboeck, G. Channel coding with multi-level/phase signals. *IEEE Trans. Inform. Theory*, IT-28 (1), 55–67, 1982.
- 13 Forney, G D Jr. et al. Efficient modulation for band-limited channels. *IEEE J. Select. Areas Commun.*, SAC-2 (5), 632–647, 1984.
- 14 Wei, L-F. Trellis-coded modulation with multidimensional constellations. *IEEE Trans. Inform. Theory*, 33 (4), 483–501, 1987.
- 15 Pietrobon, S S, Costello, D J Jr. Trellis coding with multidimensional QAM signal sets. *IEEE Trans. Inform. Theory*, 39 (2), 325–336, 1993.
- 16 Wang, F-Q, Costello, D J Jr. New rotationally invariant four-dimensional trellis codes. *IEEE Trans. Inform. Theory*, 42 (1), 291–300, 1996.
- 17 Forney, G D Jr., Ungerboeck, G. Modulation and coding for linear Gaussian channels. *IEEE Trans. Inform. Theory*, 44 (6), 2384–2415, 1998.
- 18 Le Goff, S, Glavieux, A, Berrou, C. Turbo-codes and high spectral efficiency modulation. In: *Proc. IEEE Int. Conf. Commun. (ICC'94)*, New Orleans, Louisiana, May 1994, 645–649.
- 19 Robertson, P, Wörz, T. A novel bandwidth efficient coding scheme employing turbo codes. In: *Proc. IEEE Int. Conf. Commun. (ICC'96)*, Dallas, Texas, June 1996, 962–967.
- 20 S. Benedetto, S et al. Parallel concatenated trellis-coded modulation. In: *Proc. IEEE Int. Conf. Commun. (ICC'96)*, Dallas, Texas, June 1996, 974–978.
- 21 Vucetic, B, Yuan, J. *Turbo Codes : Principles and Applications*. Norwell, MA, Kluwer, 2000.
- 22 Webb, W T, Hanzo, L. *Modern Quadrature Amplitude Modulation*. Graham Lodge, London, Pentech Press, 1994.
- 23 Alouini, M-S, Goldsmith, A J. Capacity of Nakagami multipath fading channels. In: *Proc. 47th IEEE Vehicular Technology Conference (VTC'97)*, Phoenix, Arizona, May 1997, 358–362.
- 24 Alouini, M-S, Goldsmith, A J. Adaptive M-QAM modulation over Nakagami fading channels. In: *Proc. 6th Communications Theory Mini-Conference (CTMC VI)* in conjunction with IEEE Global Communications Conference (GLOBECOM'97), Phoenix, Arizona, Nov. 1997, 218–223.
- 25 Cover, T M, Thomas, J A. *Elements of Information Theory*. New York, John Wiley, 1991.
- 26 Stüber, G L. *Principles of Mobile Communication*. Norwell, MA, Kluwer Academic Publishers, 1996.
- 27 Gradshteyn, I S, Ryzhik, I M. *Table of Integrals, Series, and Products*. San Diego, CA, Academic Press, fifth ed., 1994.
- 28 Wang, F-Q, Costello, D J Jr. Erasure-free sequential decoding of trellis codes. *IEEE Trans. Inform. Theory*, 40 (6), 1803–1817, 1994.
- 29 Couturier, S, Costello, D J Jr., Wang, F-Q. Sequential decoding with trellis shaping. *IEEE Trans. Inform. Theory*, 41 (6), 2037–2040, 1995.
- 30 Wang, F-Q, Costello, D J Jr. Sequential decoding of trellis codes at high spectral efficiencies. *IEEE Trans. Inform. Theory*, 43 (6), 2013–2019, 1997.

Breaking the Barriers of Shannon's Capacity: An Overview of MIMO Wireless Systems

DAVID GESBERT AND JABRAN AKHTAR



David Gesbert (32) holds an MSc from the Nat. Inst. for Telecommunications, Evry, France, 1993, and a PhD from Ecole Nat. Supérieure des Telecommunications, Paris, 1997. He has worked with France Telecom Research and been a postdoctoral fellow in the Information Systems Lab., Stanford University. In 1998 he took part in the founding team of Iospan Wireless Inc., San Jose, a company promoting high-speed wireless Internet access networks. In 2001 he joined the Signal Processing Group at the Univ. of Oslo as adjunct associate professor. Dr. Gesbert's research interests are in the area of high-speed wireless data / IP networks, smart antennas and MIMO, link layer and system optimization.
gesbert@ifi.uio.no



Jabran Akhtar (25) is currently a PhD student at the University of Oslo. His research interests include MIMO systems and space-time coding techniques.
jabrana@ifi.uio.no

Appearing a few years ago in a series of information theory articles published by members of the Bell Labs, multiple-input multiple-output (MIMO) systems have evolved quickly to both become one of the most popular topics among wireless communication researchers and reach a spot in today's 'hottest wireless technology' list. In this overview paper, we come back on the fundamentals of MIMO wireless systems and explain the reasons of their success, triggered mainly by the attraction of radio transmission capacities far greater than those available today. We also describe some practical transmission techniques used to signal data over MIMO links and address channel modeling issues. The challenges and limitations posed by deploying this technology in realistic propagation environment are discussed as well.

I. Introduction

Digital communications using MIMO (multiple-input multiple-output), or sometimes called "volume to volume" wireless links, has emerged as one of the most promising research areas in wireless communications. It also figures prominently on the list of hot technologies that may have a chance to resolve the bottlenecks of traffic capacity in the forthcoming high-speed broadband wireless Internet access networks (UMTS¹⁾ and beyond).

MIMO systems can be defined simply. Given an arbitrary wireless communication system, MIMO refers to a link for which the transmitting end as well as the receiving end is equipped with multiple antenna elements, as illustrated in Figure 1. The idea behind MIMO is that the signals on the transmit antennas on one end and that of the receive antennas on the other end are "combined" in such a way that the quality (Bit Error Rate) or the data rate (Bit/Sec) of the communication will be improved. MIMO systems use space-time processing techniques in that the time dimension (natural dimension of transmission signals) is completed with the spatial dimension brought by the multiple antennas. MIMO systems can be viewed as an extension of the so-called "smart antennas" [1], a popular technology for improving wireless transmission that was first invented in the 70s. However, as we

will see here, the underlying mathematical nature of MIMO environments can give performance which goes well beyond that of conventional smart antennas. Perhaps the most striking property of MIMO systems is the ability to turn multipath propagation, usually a pitfall of wireless transmission, into an advantage for increasing the user's data rate, as was first shown in groundbreaking papers by J. Foschini [2], [3].

In this paper, we attempt to explain the promise of MIMO techniques and explain the mechanisms behind it. To highlight the specifics of MIMO systems and give the necessary intuition, we illustrate the difference between MIMO and conventional smart antennas in section II. A more theoretical (information theory) standpoint is taken in part III. Practical design of MIMO solutions involves both transmission algorithms and channel modeling to measure their performance. These issues are addressed in sections IV and V respectively. Radio network level considerations to evaluate the overall benefits of MIMO setups are finally discussed in section VI.

II. MIMO Systems: More Than Smart Antennas

In the conventional wireless terminology, smart antennas refer to those signal processing techniques exploiting the data captured by multiple antenna elements located on one end of the link



Figure 1 Diagram for a MIMO wireless transmission system. The transmitter and receiver are equipped with multiple antenna elements. Coding, modulation and mapping of the signals onto the antennas may be realized jointly or separately

¹⁾ Universal Mobile Telephone Services.

only, typically at the base station (BTS) where the extra cost and space are more easily affordable. The multiple signals are combined upon transmission before launching into the channel or upon reception. The goal is to offer a more reliable communications link in the presence of adverse propagation conditions such as multipath fading and interference. A key concept in smart antennas is that of beamforming by which one increases the average signal to noise ratio (SNR) through focusing energy into desired directions. Indeed, if one estimates the response of each antenna element to a desired transmitted signal, one can optimally combine the elements with weights selected as a function of each element response. One can then maximize the average desired signal level and minimize the level of other components (noise and/or interference).

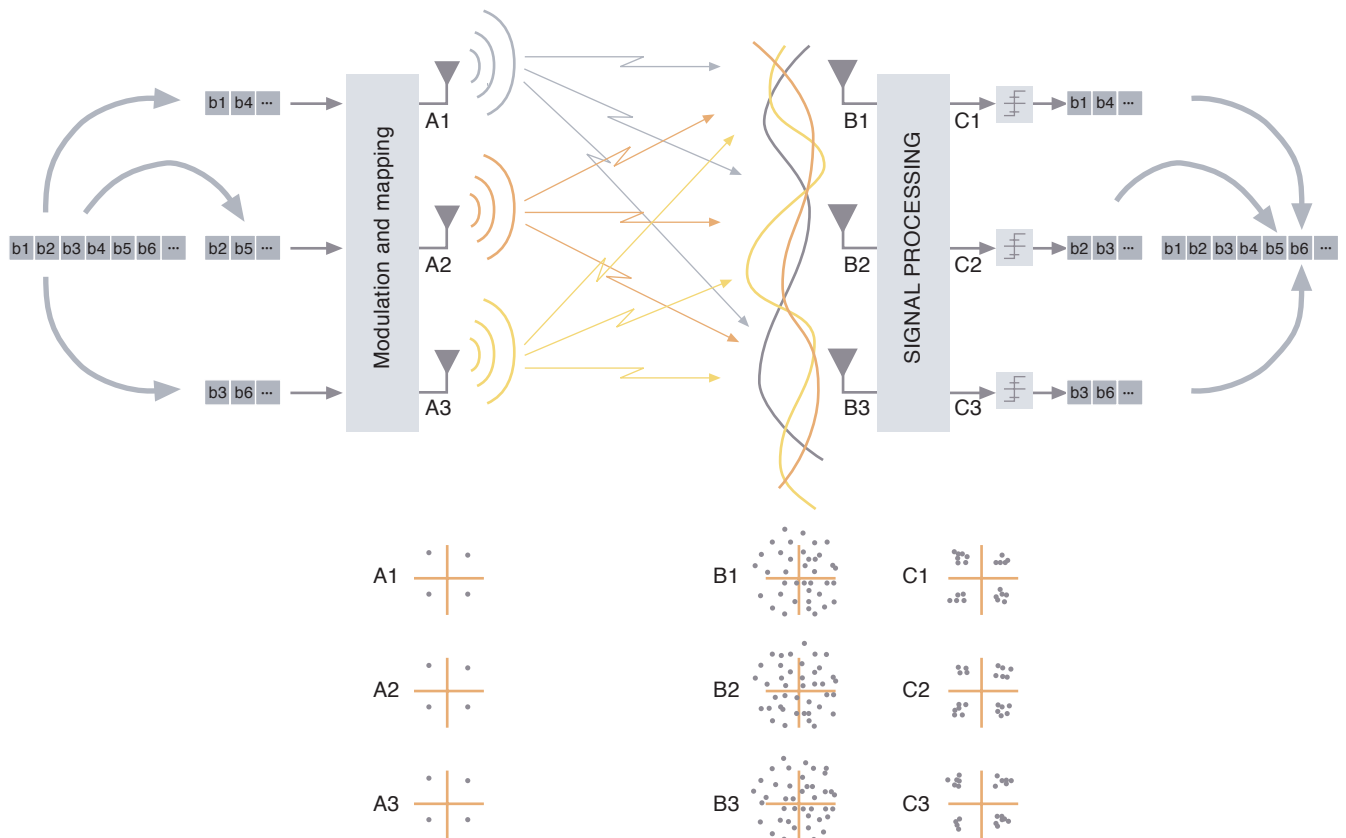
Another powerful effect of smart antennas is called *spatial diversity*. In the presence of multipath, the received power level is a random function of the user location and, at times, experiences *fading*. When using antenna arrays, the probability of losing the signal altogether vanishes exponentially with the number of decorrelated antenna elements. The *diversity order* is defined by the number of decorrelated spatial branches.

When multiple antennas are added at the subscriber's side as well as to form a MIMO link, conventional benefits of smart antennas are

retained since the optimization of the transmitting and receiving antenna elements can be carried out in a larger space. But in fact MIMO links offer advantages which go far beyond that of smart antennas [4]. Multiple antennas at both the transmitter and the receiver create a *matrix* channel (of size the number of receive antennas times the number of transmit antennas). The key advantage lies in the possibility of transmitting over several spatial *modes* of the matrix channel within the same time-frequency slot at no additional power expenditure.

While we use information theory below to demonstrate this rigorously, the best intuition is perhaps given by a simple example of a transmission algorithm over MIMO referred here as *spatial multiplexing*, which was initially described in [3], [5]. In Figure 2, a high rate bit stream (left) is decomposed into three independent bit sequences, which are then transmitted simultaneously using multiple antennas. The signals are launched and naturally mixed together into the wireless channel as they use the same frequency spectrum. At the receiver, after having identified the mixing channel matrix through training symbols, the individual bit streams are separated and estimated. This occurs in the same way, as three unknowns are resolved from a linear system of three equations. The separation is possible only if the equations are independent which can be interpreted by each antenna 'seeing' a sufficiently different channel. That is typi-

Figure 2 Basic spatial multiplexing (SM) scheme with 3 transmit and 3 receive antennas yielding three-fold improvement in spectral efficiency



cally the case in the presence of rich multipath. Finally the bits are merged together to yield the original high rate signal.

In general though, one will define the *rank* of the MIMO channel as the number of independent equations offered by the linear system mentioned above. It is also equal to the algebraic rank of the channel matrix. Clearly the rank is always both less than the number of transmit antennas and less than the number of receive antennas. In turn, the number of independent signals that one may safely transmit through the MIMO system is at most equal to the rank. In this example, the rank is assumed full (equal to three) and the system shows a spectrum efficiency gain of three. This surprising result can be demonstrated from an information theory standpoint.

III. Fundamental Limits of Wireless Transmission

Today's inspiration for research and applications of wireless MIMO systems was mostly triggered by the initial Shannon capacity results obtained independently by Bell Lab's researchers E. Telatar [6] and J. Foschini [3], further demonstrating the seminal role of information theory in telecommunications. The analysis of information theory-based channel capacity gives very useful, although idealistic, bounds on what is the maximum information transfer rate one is able to realize between two points of a communication link modeled by a given channel. Further, the analysis of theoretical capacity gives information on how the channel model or the antenna setup itself may influence the transmission rate. Finally it helps the system designer benchmark transmitter and receiver algorithm performance. Here we examine the capacity aspects of MIMO systems compared with single input single output (SISO), single input multiple output (SIMO) and multiple input single output (MISO) systems.

III.A Shannon Capacity of Wireless Channels

Given a single channel corrupted by an additive white Gaussian noise (AWGN), at a level of SNR denoted by ρ , the capacity (rate that can be achieved with no constraint on code or signaling complexity) can be written as [7]:

$$C = \log_2 (1 + \rho) \text{ Bit/Sec/Hz} \quad (1)$$

This can be interpreted by an increase of 3 dB in SNR required for each extra bit per second per Hertz. In practice, wireless channels are time-varying and subject to random fading. In this case we denote h the unit-power complex Gaussian amplitude of the channel at the instant of observation. The capacity, written as:

$$C = \log_2 (1 + \rho |h|^2) \text{ Bit/Sec/Hz} \quad (2)$$

becomes a *random* quantity, whose distribution can be computed. The cumulative distribution of this "1 x 1" case (one antenna on transmit and one on receive) is shown on the left in Figure 3. We notice that the capacity takes, at times, very small values, due to fading events.

Interesting statistics can be extracted from the random capacity related with different practical design aspects. The *average capacity* C_a , average of all occurrences of C , gives information on the average data rate offered by the link. The *outage capacity* C_o is defined as the data rate that can be guaranteed with a high level of certainty, for a reliable service:

$$\text{Prob}\{C \geq C_o\} = 99.9..9 \% \quad (3)$$

We will now see that MIMO systems affect C_a and C_o in different ways than conventional smart antennas do. In particular MIMO systems have the unique property of significantly increasing both C_a and C_o .

III.B Multiple Antennas at One End

Given a set of M antennas at the receiver (SIMO system), the channel is now composed of M distinct coefficients $\mathbf{h} = [h_0, h_1, \dots, h_{M-1}]$ where h_i is the channel amplitude from the transmitter to the i -th receive antenna. The expression for the random capacity (2) can be generalized to [3]:

$$C = \log_2 (1 + \rho \mathbf{h} \mathbf{h}^*) \text{ Bit/Sec/Hz} \quad (4)$$

where $*$ denotes the transpose conjugate. In Figure 3 we see the impact of multiple antennas on the capacity distribution with 8 and 19 antennas respectively. Both the outage area (bottom of the curve) and the average (middle) are improved. This is due to the spatial diversity which reduces fading and thanks to the higher SNR of the combined antennas. However going from 8 to 19 antennas does not give very significant improvement as spatial diversity benefits quickly level off. The increase in average capacity due to SNR improvement is also limited because the SNR is increasing inside the log function in (4). We also show the results obtained in the case of multiple transmit antennas and one receive antennas, "8 x 1" and "19 x 1" when the transmitter does not know the channel in advance (typical for a frequency duplex system). In such circumstances the outage performance is improved but not the average capacity. That is because multiple transmit antennas cannot beamform blindly.

In summary, conventional multiple antenna systems are good at improving the outage capacity performance, attributable to the spatial diversity

effect but this effect saturates with the number of antennas.

III.C Capacity of MIMO Links

We now consider a full MIMO link as in Figure 1 with respectively N transmit and M receive antennas. The channel is represented by a matrix of size $M \times N$ with random independent elements denoted by \mathbf{H} . It was shown in [3] that the capacity, still in the absence of transmit channel information, is derived from:

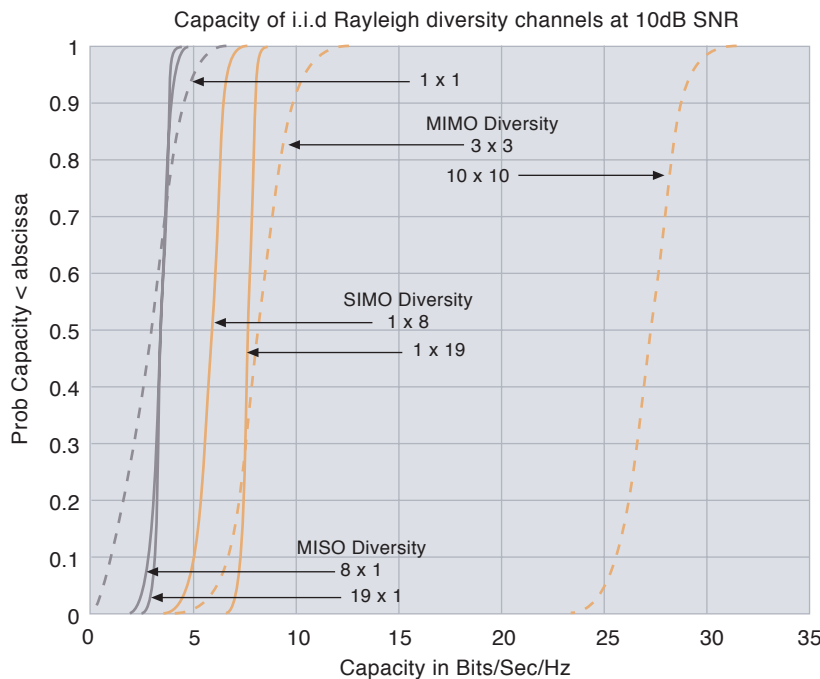
$$C = \log_2 \left[\det \left(\mathbf{I}_M + \frac{\rho}{N} \mathbf{H} \mathbf{H}^* \right) \right], \quad (5)$$

where ρ is the average SNR at any receiving antenna. In Figure 3 we have plotted the results for the 3×3 and the 10×10 case, giving the same total of 9 and 20 antennas as previously. The advantage of the MIMO case is significant, both in average and outage capacity. In fact, for a large number $M = N$ of antennas the average capacity increases linearly with M :

$$C_a \approx M \log_2 (1 + \rho) \quad (6)$$

In general the capacity will grow proportional with the smallest number of antennas $\min(N, M)$ outside and no longer inside the log function. Therefore in theory and in the case of idealized random channels, limitless capacities can be realized provided we can afford the cost and space of many antennas and RF chains. In reality the performance will be dictated by the practical transmission algorithms selected and by the physical channel characteristics.

Figure 3 Shannon capacity as function of number of TX \times RX antennas. The plots show the so-called cumulative distribution of capacity. For each curve, the bottom and the middle give indication of the outage performance and average data rate respectively



IV. Data Transmission over MIMO Systems

A usual pitfall of information theoretic analysis is that it does not reflect the performance achieved by actual transmission systems, since it is an upper bound realized by algorithms/codes with boundless complexity. The development of algorithms with reasonable performance/complexity compromise is required to realize the MIMO gains in practice. Here we give the intuition behind key transmission algorithms and compare their performance.

IV.A General Principles

Current transmission schemes over MIMO typically fall into two categories: Data rate maximization or diversity maximization schemes. The first kind focuses on improving the average capacity behavior. For example in the case of Figure 2, the objective is just to perform spatial multiplexing as we send as many independent signals as we have antennas.

More generally, however, the individual streams should be encoded jointly in order to protect transmission against errors caused by channel fading. This leads to a second kind of approach in which one tries also to minimize the outage probability.

Note that if the level of coding is increased between the transmit antennas, the amount of independence between the signals decreases. Ultimately it is possible to code the signals so that the effective data rate is back to that of a single antenna system. Effectively each transmit antenna then sees a differently encoded version of the same signal. In this case the multiple antennas are only used as a source of spatial diversity and not to increase data rate directly.

The set of schemes allowing to adjust and optimize joint encoding of multiple transmit antennas are called *space-time codes* (STC). Although STC schemes were originally revealed in [8] in the form of convolutional codes for MISO systems, the popularity of such techniques really took off with the discovery of the so-called *space-time block codes* (STBC). In contrast to convolutional codes, which require computation-hungry trellis search algorithms at the receiver, STBC can be decoded with much simpler linear operators, at little loss of performance. In the interest of space and clarity we limit ourselves to an overview of STBC below. A more detailed summary of the whole area can be found in [9].

IV.B Maximizing Diversity with Space-Time Block Codes

The field of space-time block coding was initiated by Alamouti [10] in 1998. The objective behind this work was to place two antennas at

the transmitter side and thereby provide an order two diversity advantage to a receiver with only a single antenna, with no *a priori* channel information at the transmitter. The very simple structure of Alamouti's method itself makes it a very attractive scheme that is currently being considered in UMTS standards.

The strategy behind Alamouti's code is as follows. The symbols to be transmitted are grouped in pairs. Because this scheme is a pure diversity scheme and results in no rate increase²⁾ we take two symbol durations to transmit a pair of symbols, such as s_0 and s_1 . We first transmit s_0 on the first antenna while sending s_1 simultaneously on the second one. In the next time-interval $-s_1^*$ is sent from the first antenna while s_0^* from the second one. In matrix notation, this scheme can be written as:

$$\mathbf{C} = \frac{1}{\sqrt{2}} \begin{pmatrix} s_0 & -s_1^* \\ s_1 & s_0^* \end{pmatrix}. \quad (7)$$

The rows in the code matrix \mathbf{C} denote the antennas while the columns represent the symbol period indexes. As one can observe the block of symbols s_0 and s_1 are coded across time and space, giving the name space-time block code to such designs. The normalization factor additionally ensures that the total amount of energy transmitted remains at the same level as in the case of one transmitter.

The two (narrow-band) channels from the two antennas to the receiver can be placed in a vector format as $\mathbf{h} = [h_0, h_1]$. The receiver collects observations over two time frames in a vector \mathbf{y} which can then be written as $\mathbf{y} = \mathbf{h}\mathbf{C} + \mathbf{n}$ or equivalently as $\mathbf{y}^t = \hat{\mathbf{H}}\mathbf{s} + \mathbf{n}$, where

$$\hat{\mathbf{H}} = \frac{1}{\sqrt{2}} \begin{pmatrix} h_0 & h_1 \\ h_1^* & -h_0^* \end{pmatrix}, \quad \mathbf{s} = [s_0, s_1]^T \text{ and } \mathbf{n} \text{ is}$$

the noise vector.

Because the matrices $\mathbf{C}, \hat{\mathbf{H}}$ are orthogonal by design, the symbols can be separated/decoded in a simple manner from filtering of the observed vector \mathbf{y} . Furthermore, each symbol comes with a diversity order of two exactly. Notice finally this happens despite the channel coefficients being unknown to the transmitter.

More recently some authors have tried to extend the work of Alamouti to more than two transmit antennas [11], [12]. It turns out however that in that case it is not possible to design a perfectly orthogonal code, except for real valued modulations (e.g. PAM). In the case of a general complex symbol constellation, full-rate orthogonal codes cannot be constructed. This has therefore

led to a variety of code design strategies to prolong Alamouti's work where one either sacrifices the data rate to preserve a simple decoding structure or the orthogonality of the code to retain a full data rate [13], [14], [15]. Although transmit diversity codes have mainly been designed with multiple transmit and single receive antenna in mind, the same ideas can easily be expanded towards a full MIMO setup. The Alamouti code implemented on a system with two antennas at both transmitter and receiver side will for example give a four-order diversity advantage to the user and still has a simple decoding algorithm. However, in a MIMO situation, one would not only be interested in diversity but also in increasing the data rate as shown below.

IV.C Spatial Multiplexing

Spatial multiplexing, or V-BLAST (Vertical Bell Labs Layered Space-Time) [3], [16] can be regarded as a special class of space-time block codes where streams of independent data are transmitted over different antennas, thus maximizing the average data rate over the MIMO system. One may generalize the example given in II in the following way: Assuming a block of independent data \mathbf{C} is transmitted over the $N \times M$ MIMO system, the receiver will obtain $\mathbf{Y} = \mathbf{H}\mathbf{C} + \mathbf{N}$. In order to perform symbol detection, the receiver must un-mix the channel, in one of various possible ways. Zero-forcing techniques use a straight matrix inversion, a simple approach that can also result in poor results when the matrix \mathbf{H} becomes very ill-conditioned in certain random fading events. The optimum decoding method on the other hand is known as maximum likelihood (ML) where the receiver compares all possible combinations of symbols that could have been transmitted with what is observed:

$$\hat{\mathbf{C}} = \arg \min_{\hat{\mathbf{C}}} \|\mathbf{Y} - \mathbf{H}\hat{\mathbf{C}}\| \quad (8)$$

The complexity of ML decoding is high, and even prohibitive when many antennas or high order modulations are used. Enhanced variants of this, like sphere decoding [17] have recently been proposed. Another popular decoding strategy proposed alongside V-BLAST is known as nulling and canceling which gives a reasonable tradeoff between complexity and performance. The matrix inversion process in nulling and canceling is performed in layers where one estimates a symbol, subtracts this symbol estimate from \mathbf{Y} and continues the decoding successively [3].

Straight spatial multiplexing allows for full independent usage of the antennas, however it gives

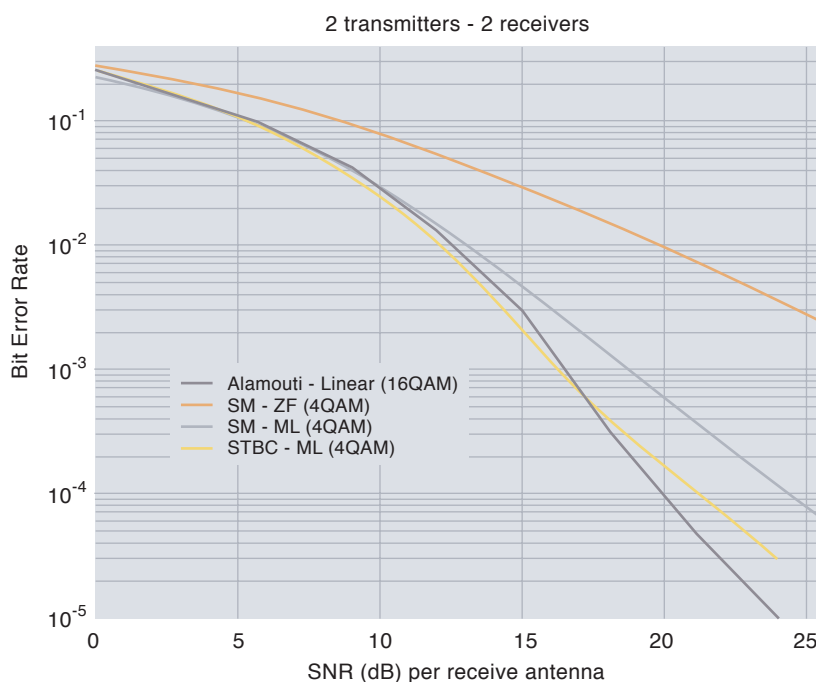
²⁾ Diversity gains can however be used to increase the order of the modulation.

limited diversity benefit and is not always the best transmission scheme for a given BER target. Coding the symbols within a block can result in additional coding and diversity gain, which can help improve the performance, even though the data rate is kept at the same level. It is also possible to sacrifice some data rate for more diversity. Methods to design such codes start from a general structure where one often assumes that a weighted linear combination of symbols may be transmitted from any given antenna at any given time. The weights themselves are selected in different fashions by using analytical tools or optimizing various cost functions [11], [18], [19], [20].

In what follows we compare four transmission strategies over a 2×2 MIMO system with ideally uncorrelated elements. All schemes result in the same spectrum efficiency but offer different BER performance.

Figure 4 shows such a plot where the BER of various approaches are compared: The Alamouti code in [7], spatial multiplexing (SM) with zero forcing (ZF) and with maximum likelihood decoding (ML), and a combined STBC spatial multiplexing scheme [20]. A 4-QAM constellation is used for the symbols except for the Alamouti code, which is simulated under 16-QAM to keep the data rate at the same level. It can be seen from the figure that spatial multiplexing with zero-forcing returns rather poor results, while the curves for other coding methods are more or less closer to each other. Coding schemes, such as Alamouti and the block code give better results than what can be achieved

Figure 4 Bit Error Rate (BER) comparisons for various transmission techniques over MIMO. All scheme results use the same transmission rate



with spatial multiplexing alone for the case of two antennas. The Alamouti curve has the best slope at high SNR because it focuses entirely on diversity (order four). At lower SNR, the scheme combining spatial multiplexing with some block coding is the best one.

It is important to note that as the number of antennas increases, the diversity effect will give diminishing returns, while the data rate gain of spatial multiplexing remains linear with the number of antennas. Therefore, for a larger number of antennas it is expected that more weight has to be put on spatial multiplexing and less on space-time coding. Interestingly, having a larger number of antennas does not need to result in a larger number of RF chains. By using antenna selection techniques (see for example [21]) it is possible to retain the benefits of a large MIMO array with just a subset of antennas being active at the same time.

V. Channel Modeling

Channel modeling has always been an important area in wireless communications and this area of research is particularly critical in the case of MIMO systems. In particular, as we have seen earlier, the promise of high MIMO capacities largely relies on decorrelation properties between antennas as well as the full-rankness of the MIMO channel matrix. The performance of MIMO algorithms such as those above can vary enormously depending on the realization or not of such properties. In particular, spatial multiplexing becomes completely inefficient if the channel has rank one. The final aim of channel modeling is therefore to get an understanding of, by the means of converting measurement data into tractable formulas, what performance can be reasonably expected from MIMO systems in practical propagation situations. The other role of channel models is to provide with the necessary tools to analyze the impact of selected antenna or propagation parameters (spacing, frequency, antenna height, etc.) onto the capacity to influence the system design in the best way. Finally, models are used to try out transmit and receive processing algorithms with more realistic simulation scenarios than those normally assumed in the literature.

V.A Theoretical Models

The original papers on MIMO capacity used an 'idealistic' channel matrix model consisting of perfectly uncorrelated (i.i.d.) random Gaussian elements. This corresponds to a rich multipath environment, yielding maximum excitation of all channel modes. It is also possible to define other types of theoretical models for the channel matrix \mathbf{H} , which are not as ideal. In particular we emphasize the *separate* roles played by antenna correlation (on transmit or on receive)

and the rank of the channel matrix. If fully correlated antennas will lead to a low rank channel, the converse is not true in general.

Let us next consider the following MIMO theoretical model classification, starting from Foschini's ideal i.i.d. model, and interpret the performance. In each case below we consider a frequency-flat channel. In the case of broadband, frequency selective channels, a different frequency-flat channel can be defined at each frequency.

- *Uncorrelated High Rank (UHR, a.k.a. i.i.d.)* model: The elements of \mathbf{H} are i.i.d. complex Gaussian.
- *Correlated Low Rank (CLR) model:* $\mathbf{H} = g_{rx} \mathbf{s}_{rx}^* \mathbf{u}_{rx} \mathbf{u}_{tx}^*$ where g_{rx} and g_{tx} are independent Gaussian coefficients (receive and transmit fading) and \mathbf{u}_{rx} and \mathbf{u}_{tx} are fixed deterministic vectors of size $M \times 1$ and $N \times 1$, respectively, and with unit modulus entries. This model is obtained when antennas are placed too close to each other or there is too little angular spread at both the transmitter and the receiver. This case yields no diversity nor multiplexing gain whatsoever, just receive array / beam-forming gain. We may also imagine the case of uncorrelated antennas at the transmitter and decorrelated at the receiver, or vice versa.
- *Uncorrelated Low Rank (ULR) (or "pin-hole" [22])* model: $\mathbf{H} = \mathbf{g}_{rx} \mathbf{g}_{tx}^*$, where \mathbf{g}_{rx} and \mathbf{g}_{tx} are independent receive and transmit fading vectors with i.i.d. complex-valued components. In this model every realization of \mathbf{H} has rank 1 despite uncorrelated transmit and receive antennas. Therefore, although diversity is present capacity must be expected to be less than in the UHR model since there is no multiplexing gain. Intuitively, in this case the diversity order is equal to $\min(M, N)$.

V.B Heuristic Models

In practice of course, the complexity of radio propagation is such that MIMO channels will not fall completely in either of the theoretical cases described above. Antenna correlation and matrix rank are influenced by as many parameters as the antenna spacing, the antenna height, the presence and disposition of local and remote scatterers, the degree of line of sight and more. Figure 5 depicts a general setting for MIMO propagation. The goal of heuristic models is to display a wide range of MIMO channel behaviors through the use of as few relevant parameters as possible with as much realism as possible.

A good model shall give us answers to the following problems: What is the typical capacity of an outdoor or indoor MIMO channel? What are

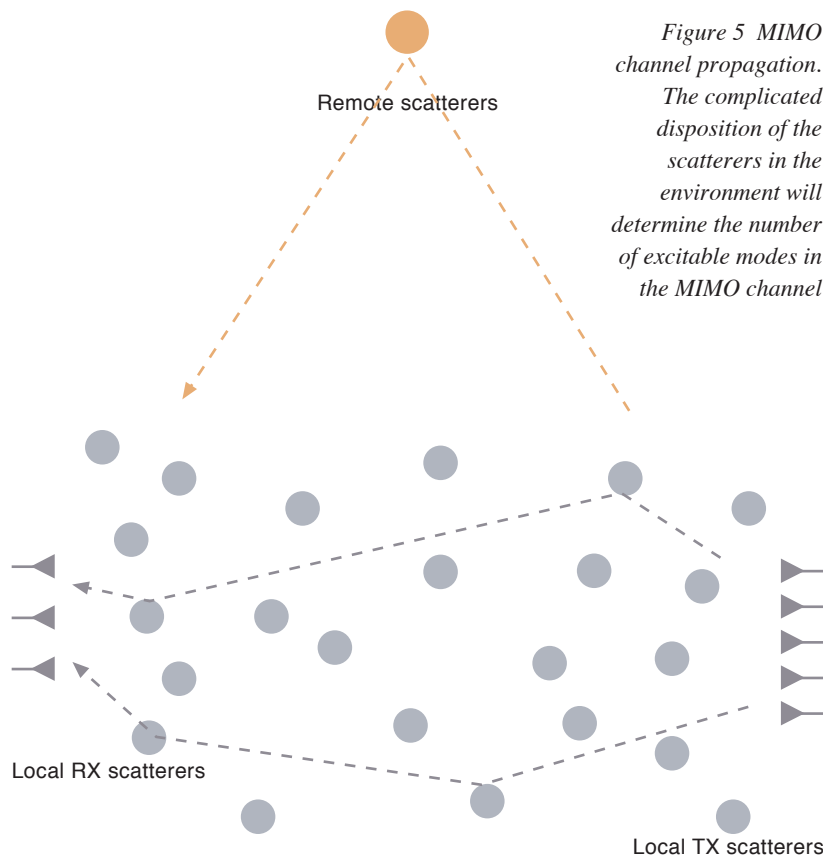


Figure 5 MIMO channel propagation. The complicated disposition of the scatterers in the environment will determine the number of excitable modes in the MIMO channel

the key parameters governing capacity? Under what simple conditions do we get a full rank channel? If possible the model parameters should be controllable (such as antenna spacing) or measurable (such as angular spread of multipath [23], [24], which is not always easy to achieve.

The literature on these problems is still very scarce. For the line-of-sight (LOS) case it has only been shown how very specific arrangements of the antenna arrays at the transmitter and the receiver can maximize the orthogonality between antenna signatures and produce maximum capacity as reported in [25]. But, in a general situation with fading, which is the true promising case, this work is not applicable.

In the presence of fading, the first step in increasing the model's realism consists in taking into account the correlation of antennas at either the transmit or receive side. The correlation can be modeled to be inversely proportional to the angular spread of the arriving/departing multipath. The experience suggests that higher correlation at the BTS side can be expected because the BTS antenna is usually higher above the clutter, causing reduced angular spread. In contrast the subscriber's antenna will be buried in the clutter (if installed at street level) and will experience more multipath angle spread, hence less correlation for the same spacing. The way

MIMO models can take correlation into account is similar to how usual smart antenna channel models do it. The channel matrix is pre- (or post-) multiplied by a correlation matrix controlling the antenna correlation as function of the path angles, the spacing and the wavelength. For example, for a MIMO channel with correlated receive antennas, we have:

$$\mathbf{H} = \mathbf{R}_{\theta_r, d_r}^{1/2} \mathbf{H}_0 \quad (9)$$

where \mathbf{H}_0 is an ideal i.i.d. MIMO channel matrix and $\mathbf{R}_{\theta_r, d_r}$ is the $M \times M$ correlation matrix. θ_r is the receive angle spread and d_r is the receive antenna spacing. Different assumptions on the statistics of the paths' directions of arrival (DOA) will yield different expressions for $\mathbf{R}_{\theta_r, d_r}$ [26], [27], [28]. For uniformly distributed DOAs, we find [27], [26]

$$\left[\mathbf{R}_{\theta_r, d_r} \right]_{m,k} = \frac{1}{S} \sum_{i=\frac{S-1}{2}}^{i=\frac{S-1}{2}} e^{-2\pi j(k-m)d_r \cos\left(\frac{\pi}{2} + \theta_{r,i}\right)} \quad (10)$$

where S (assumed odd) is the number of paths with corresponding DOAs $\theta_{r,i}$. For "large" values of the angle spread and/or antenna spacing,

$\mathbf{R}_{\theta_r, d_r}$ will converge to the identity matrix,

which gives uncorrelated fading. For "small" values of θ_r, d_r , the correlation matrix becomes rank deficient (eventually rank one) causing fully correlated fading. The impact of the correlation on the capacity was analyzed in several papers, including [29]. Note that it is possible to generalize this model to include correlation on both sides by using two distinct correlation matrices:

$$\mathbf{H} = \mathbf{R}_{\theta_r, d_r}^{1/2} \mathbf{H}_0 \mathbf{R}_{\theta_t, d_t}^{1/2} \quad (11)$$

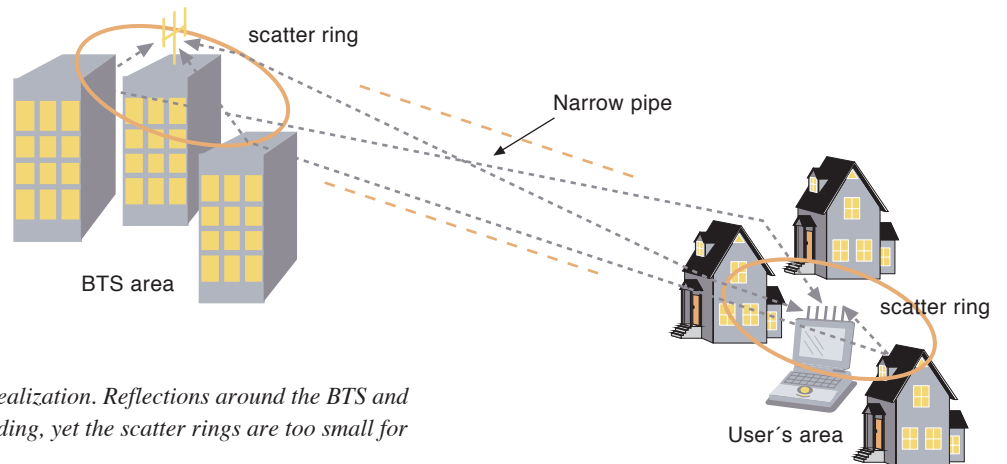


Figure 6 An example of pinhole realization. Reflections around the BTS and subscribers cause uncorrelated fading, yet the scatter rings are too small for the rank to build up

V.B.1 Impact of Scattering Radius

One limitation of simple models like the one in (11) is that it implies that rank loss of \mathbf{H} can only come from rank loss in $\mathbf{R}_{\theta_r, d_r}$ or in

$\mathbf{R}_{\theta_t, d_t}$, i.e. a high correlation between the

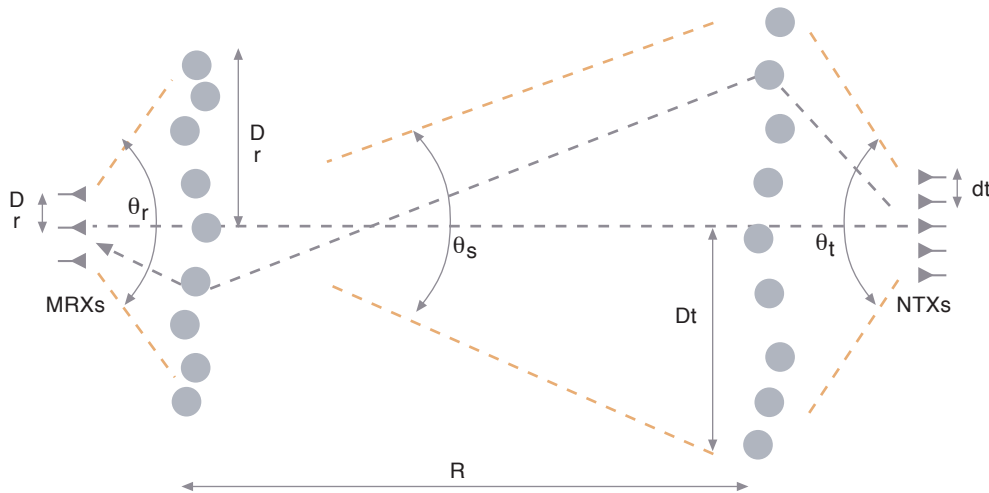
antennas. However as suggested by the theoretical model "ULR" above, it may not always be so. In practice such a situation can arise where there is significant local scattering around both the BTS and the subscriber's antenna and still only a low rank is realized by the channel matrix. That may happen because the energy travels through a narrow "pipe", i.e. if the scattering radius around the transmitter and receiver is small compared to the traveling distance. This is depicted in Figure 6. This situation is referred to as *pinhole* or *keyhole* channel in the literature [22], [30].

In order to describe the pinhole situation more, so-called *double scattering* models are developed that take into account the impact of the scattering radius at the transmitter and at the receiver. The model is based on a simplified version of Figure 5 shown in Figure 7 where only local scatterers contributing to the total aperture of the antenna as seen by the other end are considered. The model can be written as [22]:

$$\mathbf{H} = \frac{1}{\sqrt{S}} \mathbf{R}_{\theta_r, d_r}^{1/2} \mathbf{H}_{0,r} \mathbf{R}_{\theta_s, 2D_r/S}^{1/2} \mathbf{H}_{0,t} \mathbf{R}_{\theta_t, d_t}^{1/2}, \quad (12)$$

where the presence of two (instead of one) i.i.d. random matrices $\mathbf{H}_{0,t}$ and $\mathbf{H}_{0,r}$ accounts for the double scattering effect. The matrix $\mathbf{R}_{\theta_s, 2D_r/S}$ dictates the correlation between scattering elements, considered as *virtual* receive antennas with virtual aperture $2D_r$. When the virtual aperture is small, either on transmit or receive, the rank of the overall MIMO channel will fall regardless of whether the actual antennas are correlated or not.

Figure 7 Double scattering MIMO channel model



V.C Broadband Channels

In broadband applications the channel experiences frequency selective fading. In this case the channel model can be written as $\mathbf{H}(f)$ where a new MIMO matrix is obtained at each frequency/sub-band. This type of model is of interest in the case of orthogonal frequency division multiplexing (OFDM) modulation with MIMO. It was shown that the MIMO capacity actually benefits from the frequency selectivity because the additional paths that contribute to the selectivity will also contribute to a greater overall angular spread and therefore improve the average rank of the MIMO channel across frequencies [31].

V.D Measured Channels

In order to validate the models as well as to foster the acceptance of MIMO systems into wireless standards, a number of MIMO measurement campaigns have been launched in the last two years, mainly led by Lucent and ATT Labs and by various smaller institutions or companies such as Iospan wireless in California. More recently Telenor R&D put together its own MIMO measurement capability.

Samples of analysis for UMTS type scenarios can be found in [32], [33], [34]. Measurements conducted at 2.5 GHz for broadband wireless access applications can be found in [35]. So far, the results reported largely confirm the high level of dormant capacity of MIMO arrays, at least in urban or suburban environments. Indoor scenarios lead to even better results due to a very rich multipath structure. Eigenvalues analyses reveal that a large number of the modes of MIMO channels can be exploited to transmit data. Which particular combination of spatial multiplexing and space time coding will lead to the best performance complexity trade-off over such channels remains however an area of active research.

VI. System Level Issues

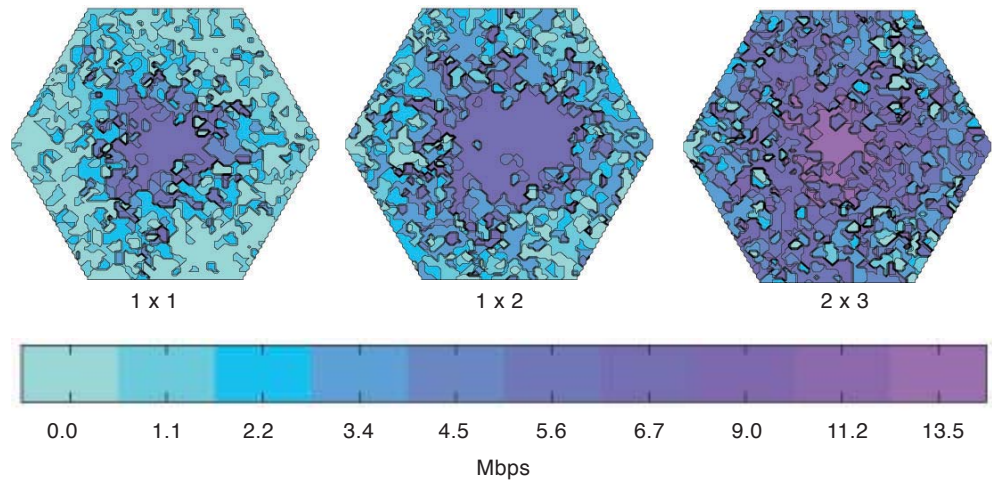
VI.A Optimum Use of Multiple Antennas

Multiple antenna techniques are not new in commercial wireless networks. Spatial diversity systems, using two or three antenna elements, co- or cross-polarized, have been in use since the early stages of mobile network deployments. More recently, beamforming-type BTS products equipped with five to ten or more antennas have been offered on the market. These products are using diversity to improve the link budget and the beamforming capability to extend the cell range or help in load balancing.

Beyond the information theory aspects addressed earlier, there are significant network-level differences between the beamforming approach and the MIMO approach to using multiple antennas.

While beamforming systems tend to use a larger number of closely spaced antennas, MIMO will operate with typically fewer antennas (although the only true constraint is at the subscriber side rather than at the BTS side). Furthermore the MIMO antennas will use as much space as can be afforded to try and realize decorrelation between the elements while the directional-based beamforming operation imposes stringent limits on spacing. Also most MIMO algorithms focus on diversity or data rate maximization rather than just increasing the average SNR at the receiver or reducing interference. Finally, beamforming systems thrive in near line of sight environments because the beams can be more easily optimized to match one or two multipaths than a hundred of them. In contrast, MIMO systems turn rich multipath into an advantage and lose their multiplexing benefits in line of sight cases.

Figure 8 User rates in 2 MHz FDD channels in a fixed wireless access system. The plots show the relative gains between various number of antennas at transmitter \times receiver (SISO, SIMO, MIMO)



Because of these differences, the optimal way of using multiple antenna systems, at least at the BTS, is likely to depend on the situation. The search for compromising solutions, in which the degrees of freedom offered by the multiple antennas are best used at each location, is an active area of work. A key to this problem resides in *adaptive* techniques, which through the tracking of environment/propagation characteristics are able to pick the right solution at all times.

VI.B MIMO in Broadband Internet Access

One unfavorable aspect of MIMO systems, when compared with traditional smart antennas, lies in the increased cost and size of the subscriber's equipment. Although a sensible design can extract significant gains with just two or three antennas at the user's side, it may already prove too much for simple mobile phone devices. Instead wireless LAN modems, PDAs and other high speed wireless Internet access, fixed or mobile, devices constitute the real opportunity for MIMO because of less stringent size and algorithmic complexity limitations. In Figure 8 we show the data rates achieved by a fixed broadband wireless access system with 2×3 MIMO. The realized user's data rates are color coded from 0 to 13.5 Mb/s in a 2 MHz RF channel³⁾, function of the user's location. The access point is located in the middle of an idealized hexagonal cell. Detailed assumptions can be found in [36]. The figure illustrates the advantages over a system with just one transmit antenna and one or two receive antennas. Current studies demonstrating the system level advantages of MIMO in wireless Internet access focus mainly on performance. While very promising,

the evaluation of overall benefits of MIMO systems, taking into account deployment and cost constraints, is still in progress.

VII. Conclusions

This paper reviews the major features of MIMO links for use in future wireless networks. Information theory reveals the great capacity gains which can be realized from MIMO. Whether we achieve this fully or at least partially in practice depends on a sensible design of transmit and receive signal processing algorithms. More progress in channel modeling will also be needed. In particular upcoming performance measurements in specific deployment conditions will be key to evaluate precisely the overall benefits of MIMO systems in real-world wireless systems scenarios such as UMTS.

References

- 1 Paulraj, A, Papadias, C B. Space-time processing for wireless communications. *IEEE Signal Proc. Mag.*, 14, 49–83, 1997.
- 2 Foschini, G J, Gans, M J. On limits of wireless communications in a fading environment when using multiple antennas. *Wireless Personal Communications*, 6, 311–335, 1998.
- 3 Foschini, G J. Layered space-time architecture for wireless communication. *Bell Labs Technical Journal*, 1, 41–59, 1996.
- 4 Sheikh, K et al. Smart antennas for broadband wireless access. *IEEE Communications Magazine*, Nov 1999.
- 5 Paulraj, A J, Kailath, T. *Increasing capacity in wireless broadcast systems using dis-*

³⁾ A user gets zero if the link quality does not satisfy the target BER.

- tributed transmission/directional reception. U.S. Patent, 1994. (No. 5,345,599.)
- 6 Telatar, I E. Capacity of multi-antenna Gaussian channels. *Bell Labs Technical Memorandum*, 1995.
 - 7 Proakis, J G. *Digital Communications*. New York, McGraw-Hill, 1989.
 - 8 Tarokh, V, Seshadri, N, Calderbank, A R. Space-time codes for high data rate wireless communication: Performance criterion and code construction. *IEEE Trans. Inf. Theory*, 44, 744–765, 1998.
 - 9 Naguib, A, Seshadri, N, Calderbank, R. Increasing data rate over wireless channels. *IEEE Signal Processing Magazine*, May 2000.
 - 10 Alamouti, S M. A simple transmit diversity technique for wireless communications. *IEEE Journal on Selected Areas in Communications*, 16, 1451–1458, 1998.
 - 11 Tarokh, V, Jafarkhani, H, Calderbank, A R. Space-time block codes for wireless communications: Performance results. *IEEE Journal on Selected Areas in Communications*, 17, 1999.
 - 12 Ganesan, G, Stoica, P. Space-time diversity using orthogonal and amicable orthogonal designs. *Wireless Personal Communications*, 18, 165–178, 2001.
 - 13 Jafarkhani, H. A quasi orthogonal space-time block code. *IEEE Trans. Comm.*, 49, 1–4, 2001.
 - 14 Tirkkonen, O, Boariu, A, Hottinen, A. Minimal non-orthogonality rate 1 space-time block code for 3+ tx antennas. In: *Proc. IEEE Int. Symp. Spread Spectrum Technology*, 2000.
 - 15 Tarokh, V, Jafarkhani, H, Calderbank, A R. Space-time block codes from orthogonal designs. *IEEE Trans. Inf. Theory*, 45, 1456–1467, 1999.
 - 16 Golden, G D et al. Detection algorithm and initial laboratory results using the V-BLAST space-time communication architecture. *Electronics Letters*, 35, 1, 14–15, 1999.
 - 17 Damen, M O, Chkeif, A, Belfiore, J C. Lattice codes decoder for space-time codes. *IEEE Communications Letters*, 4, 161–163, 2000.
 - 18 Hassibi, B, Hochwald, B. High rates codes that are linear in space and time. Submitted to *IEEE Trans. On Information Theory*, 2000.
 - 19 Sandhu, S, A. Paulraj. Unified design of linear space-time block-codes. *IEEE Globecom Conference*, 2001.
 - 20 Damen, M O, Tewfik, A, Belfiore, J C. A construction of a space-time code based on number theory. *IEEE Trans. On Information Theory*, March 2002.
 - 21 Molisch, A, Winters, M Z W J, Paulraj, A. Capacity of mimo systems with antenna selection. In: *IEEE Intern. Conf. On Communications*, 570–574, 2001.
 - 22 Gesbert, D et al. Outdoor mimo wireless channels: Models and performance prediction. *IEEE Trans. Communications*, 2002. To appear.
 - 23 Pedersen, K I, Mogensen, P E, Fleury, B. A stochastic model of the temporal and azimuthal dispersion seen at the base station in outdoor propagation environments. *IEEE Trans. On Vehicular Technology*, 49, 2000.
 - 24 Rossi, J P, Barbot, J P, Levy, A. Theory and measurements of the angle of arrival and time delay of uhf radiowaves using a ring array. *IEEE Trans. On Antennas and Propagation*, May 1997.
 - 25 Driessen, P, Foschini, J. On the capacity formula for multiple input multiple output wireless channels: a geometric interpretation. *IEEE Trans. Comm.*, 173–176, Feb 1999.
 - 26 Ertel, R B et al. Overview of spatial channel models for antenna array communication systems. *IEEE Personal Communications*, 10–22, Feb 1998.
 - 27 Asztély, D. *On antenna arrays immobile communication systems: Fast fading and GSM base station receiver algorithms*. Royal Institute of Technology, Stockholm, Sweden, March 1996. (Tech. Rep. IR-S3-SB-9611.)
 - 28 Fuhl, J, Molisch, A F, Bonek, E. Unified channel model for mobile radio systems with smart antennas. *IEE Proc.-Radar, Sonar Navig.*, 145, 32–41, 1998.
 - 29 Shiu, D et al. Fading correlation and its effect on the capacity of multi-element antenna systems. *IEEE Trans. Comm.*, March 2000.

- 30 Chizhik, D, Foschini, G, Valenzuela, R A. Capacities of multi-element transmit and receive antennas: Correlations and keyholes. *Electronic Letters*, 1099–11, 2000.
- 31 Bölcskey, H, Gesbert, D, Paulraj, A J. On the capacity of wireless systems employing OFDM-based spatial multiplexing. *IEEE Trans. Comm.*, 2002. To appear.
- 32 Martin, C C, Winters, J, Sollenberger, N. Multiple input multiple output (mimo) radio channel measurements. In: *IEEE Vehicular Technology Conference*, Boston (MA), 2000.
- 33 Ling, J et al. Multiple transmitter multiple receiver capacity survey in Manhattan. *Electronic Letters*, 37, Aug 2001.
- 34 Buehrer, R et al. Spatial channel models and measurements for imt-2000 systems. In: *Proc. IEEE Vehicular Technology Conference*, May 2001.
- 35 Pitschaiah, S et al. Modeling of multiple-input multiple-output (mimo) radio channel based on outdoor measurements conducted at 2.5 GHz for fixed bwa applications. In: *Proc. International Conference on Communications*, 2002.
- 36 Gesbert, D et al. Technologies and performance for non line-of-sight broadband wireless access networks. *IEEE Communications Magazine*, April 2002.

An Introduction to Turbo Codes and Iterative Decoding

ØYVIND YTREHUS



Øyvind Ytrehus (42) is professor and currently the Department Chair at the Department of Informatics, University of Bergen. His research interests include error correcting properties and decoding complexity of algebraic codes, convolutional codes, turbo codes, and codes based on graphs; applications of coding theory in communication and storage; and the interaction between coding theory and cryptology.

oyvind@ii.uib.no

The discovery of turbo codes by Berrou et. al. [11] in 1993 revolutionized the theory of error-correcting codes. The purpose of this paper is

- to provide an introduction to turbo codes;
- to describe the weight distribution properties of turbo codes that allow low error probability, even when the underlying communication channel is poor;
- to explain the low complexity decoding algorithms that make turbo codes so attractive;
- to suggest how to select essential components of the turbo construction, such as interleavers and constituent codes;
- to mention that turbo coding can be used in practical situations, for example in a coded modulation scheme; and finally
- to point out the limitations of turbo codes that, so far, restrict their use in some applications.

I. Introduction

Consider the problem of sending a block \mathbf{u} of K bits over a noisy channel. With some nonzero probability, these bits will be corrupted by the noise. To combat these effects, the information block is almost always encoded with an *error-correcting code*.

An error-correcting encoder works by adding $N - K$ extra *parity-check* bits to the information block, to produce a *codeword* of N bits. The *code* is the set of codewords that arises when \mathbf{u} ranges over all possible information blocks, and the *code rate* is the ratio $R = K/N$. The codeword is transmitted over the channel. At the receiving end, a *decoder* produces an estimate $\hat{\mathbf{u}}$, based on the received message, of \mathbf{u} . If $\hat{\mathbf{u}} \neq \mathbf{u}$, we have a *decoding error*, see Figure 1, from [1]¹⁾. The parity-check bits are selected with the aim to minimize the probability of decoding error.

In his seminal paper [2], Shannon introduced the notion of the *channel capacity* C of a communication channel. He proved the following by a non-constructive argument: For a given communication channel and for any arbitrarily small ϵ , provided the code length N is sufficiently large there exists a code of any rate $R < C$ with the property that the decoding error probability with optimum decoding is less than ϵ . Conversely, for rates larger than the capacity, error free transmission is not possible.

During the last half of the 20th century, intense research efforts were devoted towards designing practical error correcting codes with a performance approaching Shannon's predictions. However, this goal turned out to be difficult to achieve, even though powerful algebraic constructions were devised [5], [8]. The discovery of turbo codes by Berrou et al. [11] in 1993 represented a major breakthrough. Recent developments in this area have produced implementable codes with performance very close to Shannon's bounds [36].

Figure 2 shows the relationship between the signal-to-noise ratio (SNR) and the bit error rate (BER) for an additive white Gaussian noise (AWGN) channel. The SNR, expressed in dB, is

$$SNR = 10 \log_{10} \frac{E_b}{N_0}, \quad (1)$$

where E_b is the average received signal energy per information bit, and N_0 is the single sided power spectral density of the Gaussian noise.²⁾ For a given code rate, Shannon's results imply that there is a threshold SNR, which is referred to as the *Shannon bound*, below which error free transmission is impossible. For a continuous input AWGN channel, the Shannon bound in dB is given as

$$SNR_C = 10 \log_{10} \frac{2^{2R} - 1}{2R} \quad (2)$$

¹⁾ For non-Norwegian readers: The story "God dag mann! – Økseskaft" is about a hearing-impaired ferryman who receives a call from the local policeman. Prior to the visit, the ferryman anticipates the policeman's questions. However, as the signal-to-noise level of the actual communication channel is too low, this results in an absurd conversation.

²⁾ Throughout this paper, commonly known facts and results will be presented without explicit references. Explanations and derivations of these results can be found in some of the monographs [8], [10], [17], [24], [29].



Figure 1 “God dag mann! – Økseskaft!”: Example of decoding error

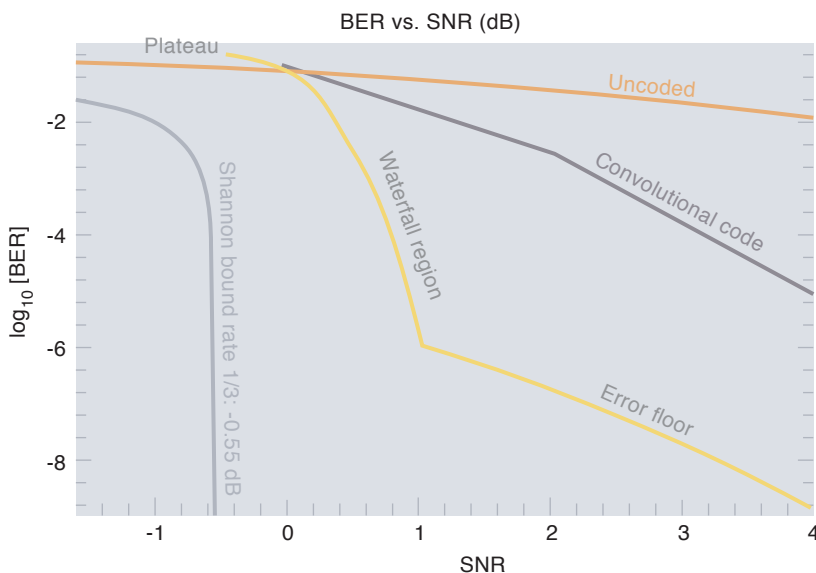


Figure 2 The connection between the SNR and the BER. The Shannon bound of (1) is shown at approximately -0.55 dB for rate $R=1/3$. Also shown are the BER curves for uncoded transmission (of rate 1), a 64-state rate $1/3$ convolutional code, and a rate $1/3$ turbo code with a simple random interleaver of information length $N = 1000$. In this example, the blue and the yellow curves intersect at an SNR of about 14 dB and a BER of about 10^{-55} . However, the turbo code can easily be improved by changing the interleaver

The BER performance of a typical turbo code used with the *iterative turbo decoding algorithm* (to be discussed in Section III) is shown in Figure 2. The proximity to the rate-specific Shannon bound varies with the information length N , with the choice of constituent encoders and interleaver, and with details of implementation of the decoding algorithms. However, over the range of such varying parameters, the curves display the characteristics as shown: There is a *plateau* region at very low SNR, where there is little or no improvement over uncoded transmission; followed by the *waterfall* region, where the BER drops off rapidly with increasing SNR. Finally there is an *error floor* at higher SNR, where the BER mainly depends on the probability of decoding to a few *most likely* error vectors.

Berrou et al.’s turbo codes are *parallel concatenated codes*, generated by an encoder as shown in Figure 3. The encoder accepts an information block $\mathbf{u} = (u_0, \dots, u_{K-1})$ consisting of K bits. It will be convenient sometimes to consider \mathbf{u} as a sequence $u(D) = \sum_{j=0}^{K-1} u_j D^j$ (where the variable D can be thought of as just a placeholder). The encoder is *systematic*, meaning that the information block \mathbf{u} is visible as an explicit part of the codeword. The rest of the codeword consists of parity check symbols from the two *constituent encoders*, encoder A and encoder B. In Berrou et al.’s model, which we will consider throughout this paper except for the generalizations in Section V, encoders A and B are identical recursive convolutional encoders. This means that the first parity sequence $c_A(D)$ is obtained from the information block as $c_A(D) = u(D)f(D)/g(D)$, for some fixed binary polynomials $f(D)$ and $g(D)$. The second parity sequence is obtained from the information block as $c_B(D) = \pi(u(D))f(D)/g(D)$, where $\pi(u(D))$ is the sequence resulting from permuting the coordinates of $u(D)$ according to an *interleaver* map π . In some cases, $2v$ extra bits are appended to the input sequences in order to terminate the constituent trellises (see Section III and [24], [29]). This is not shown in Figure 3. Finally, some of the encoded bits (usually some of the parity check bits) are punctured (deleted) according to a puncturing pattern P , leaving a total of N encoded bits. Hence the turbo code is completely specified by the two constituent encoders, by the interleaver π , by the puncturing pattern P , and by the termination rules.

In Section II we investigate the properties of the weight distribution that explain the characteristic error curves as shown in Figure 1. Section III deals with the turbo decoding algorithm. Selection of essential system components is considered in Section IV. Finally, Section V discusses

variations, generalizations, limitations, and applications.

In a short paper about a topic that in a few years has grown into a major research area, there is no space for entering into deeper discussion on the finer points. The reference list contains pointers to the literature, for those who want to pursue this subject.

II. Weight Distribution Properties: Performance at the Error Floor

When the BER is not too small, it can be accurately approximated using computer simulation. However, accurate simulation results require the observation of hundreds of error events (see any textbook in statistics or computer simulation), which are particularly hard to obtain for long turbo codes with low error floors.

The Hamming weight of a binary vector is defined as the number of nonzero positions in the vector. It is well known (see for example [10]) that, *under the assumption that maximum likelihood decoding is used*, the Frame error rate (FER; the probability that a given frame contains an error) of any error correcting code can be approximated for high SNRs by the union bound,

$$FER \leq \sum_{w=d}^N A_w Q\left(\sqrt{2wR \cdot SNR}\right), \quad (3)$$

where A_w is the number of codewords of Hamming weight w , d is the minimum distance of the code = the minimum nonzero w for which $A_w > 0$, and $Q()$ is the complementary error function,

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-t^2/2} dt. \quad (4)$$

The set of numbers $\{A_w | w = 0, \dots, N\}$ is called the *weight distribution* or the *weight spectrum* of the code. It is sometimes convenient to refer to the *weight enumerator function* $A(X) = \sum_{w=0}^N A_w X^w$. It is also often convenient to consider the *conditional input-redundancy weight enumerator function* $A_i(Z) = \sum_{z=0}^{N-K} A_{i,z} Z^z$, [14], where $A_{i,z}$ is the number of codewords of information weight i and parity check weight z . Note that

$$\begin{aligned} \sum_{i=0}^K Z^i A_i(Z) &= \sum_{i=0}^K Z^i \sum_{z=0}^{N-K} A_{i,z} Z^z \\ &= \sum_{i=0}^K \sum_{z=0}^{N-K} A_{i,z} Z^{i+z} \\ &= \sum_{w=0}^N Z^w \sum_{i=0}^K A_{i,w-i} \\ &= A(Z). \end{aligned} \quad (5)$$

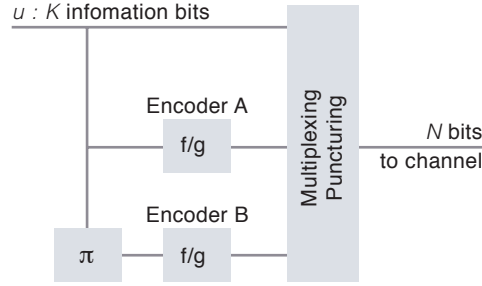


Figure 3 Turbo encoder. Trellis termination is not shown

Similarly,

$$BER \leq \sum_{w=d}^N \frac{B_w}{K} Q\left(\sqrt{2wR \cdot SNR}\right), \quad (6)$$

where $B_w = \sum_{i=1}^w A_{i,w-i}$ is the total information weight of all codewords of weight w .

The union bounds can be refined. These refinements lead to sharper bounds that also allow us to restrict the summations of equations (3) and (6) to the first few terms:

$$FER \leq \sum_{w=d}^{d+m} A_w Q\left(\sqrt{2wR \cdot SNR}\right) + \text{something that diminishes at high SNR}; \quad (7)$$

$$BER \leq \sum_{w=d}^{d+m} \frac{B_w}{K} Q\left(\sqrt{2wR \cdot SNR}\right) + \text{something that diminishes at high SNR}; \quad (8)$$

where m is some small integer. Thus in the error floor region, instead of performing a costly computer simulation, we would like to determine the first few terms of the weight distribution. For very large SNR, the code's minimum distance d_C and the total information weight of weight- d_C codewords, B_{d_C} , determine the BER quite precisely.

Benedetto and Montorsi [14] considered the case of a *uniform interleaver*; i.e. an idealized random interleaver. Under this assumption, they showed that the weight enumerator of the (basic, unpunctured) turbo code can be approximated by

$$A_i(Z) \approx \frac{A_i^A(Z) A_i^B(Z)}{\binom{K}{i}} \quad (9)$$

where $A_i^A(Z)$ and $A_i^B(Z)$, respectively, are the conditional input-redundancy weight enumerator function of the constituent codes.

These results explain the behaviour of turbo codes with a random interleaver, as observed and demonstrated by Berrou et al.:

- The actual minimum distance of a turbo code is not impressive compared to the minimum distance d_C of an arbitrary “classical” error correcting code of similar length and rate. This explains why, for a given SNR in the error floor region, the error floor is flatter for the turbo code than for the “classical” code.
- The number of codewords of moderately low weight (say, of weights not much larger than d_C) is extremely small compared to the “classical” code. This latter phenomenon is usually referred to as *spectral thinning* [14], and explains why the error floor of the turbo code is lower (for moderately low SNRs) than for the “classical” code.
- For random interleavers, the expected minimum distance grows slowly, and the expected number of codewords of low weight is slowly reduced, with the interleaver length K .

An additional observation [14] that arises from (9) is that most low weight codewords are associated with input vectors of weight 2. This observation actually makes (9) obsolete, since it motivates the design of better, non-random interleavers. Interleaver design is discussed in Section IV. For now we conclude that, since (9) does not apply to nonrandom interleavers, we need another way to determine the initial part of the weight distribution for an arbitrary interleaver.

II.A Algorithms for Determination of the Weight Distribution

In this section we consider algorithms for determining the number of codewords of weight $\leq w_{\max}$ in a particular turbo code with fixed interleaver and constituent codes.

An upper bound on the minimum distance can be achieved by considering only input vectors of low Hamming weight, say of weights 1, 2 or 3. This method can be applied for very large codes, but there is no guarantee that no lower weight codewords exist with higher weight input vectors.

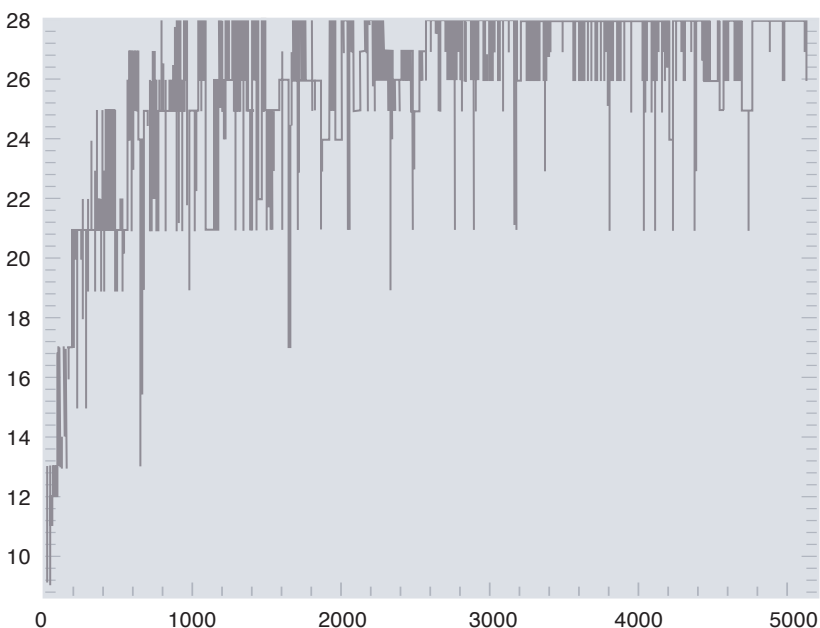
Breiling and Huber [31], refining the previous method, used pre-processing to obtain a list of all input vectors corresponding to constituent codewords of weight “almost” w_{\max} . The algorithm proceeds to attempt to combine input vectors to the two constituent encoders. The algorithm works well for small values of w_{\max} , but is prohibitively complex for larger w_{\max} .

Another approach was followed by Benedetto et al. in [37], developed further by Rosnes and Ytrehus [43]. The idea is to set up a search tree containing all possible input vectors. Each node in the search tree is a *constraint set*, which determines the value of a subset of the input vector positions. This constraint set specifies a subcode of each constituent code. The algorithm evaluates the minimum distance of each subcode, and thereby also a lower bound t on the minimum distance of the corresponding subset of the turbo code. If t exceeds w_{\max} , the constraint set can be discarded from the search tree. Otherwise, the search tree is expanded with two new nodes, expanding the current constraint set in two new directions. See details in [37]. The efficiency of the evaluation methods is discussed and improved in [43], facilitating the analysis of larger codes and larger values of w_{\max} .

The algorithm in [43] was used to determine the minimum distance of all UMTS turbo codes. The results are shown in Figure 4. As a comment on the first class of algorithms discussed in this section, it can be noted that in 351 of the 5075 cases considered in Figure 4, the minimum distance codewords are actually caused by weight-9 inputs.

As an example of the minimum distance’s impact on the BER, we also show in Figure 5 results for the UMTS codes of information length $K = 5114$. This code has a codeword of length 26. We also found another turbo code, using the same constituent code but with an interleaver generated by a technique similar to the one described in [42], with minimum distance 36. Note that at target bit error rates larger

Figure 4 The minimum distance of UMTS turbo codes of information length K ranging from 40 to 5114



than 10^{-6} , the performance of the two turbo codes is almost identical. However, at a target BER of 10^{-8} , the code with the larger minimum distance has a relative coding gain of almost 1 dB, i.e. it requires 1 dB less SNR to achieve the same BER. The asymptotic coding gain difference is approximately 1.2 dB.

II.B Distance Bounds

Upper bounds on the minimum distance of turbo codes with arbitrary interleavers were derived by Breiling and Huber [40]. They consider only codewords of input weight two and four of the types shown in Figure 6, (a) and (b), respectively. An input vector containing a subvector of $10^{mp-1}1$ (a one followed by $mp-1$ zeros followed by a one), where m is a positive integer and p is an integer called the *period* of the encoder denominator polynomial $g(D)$, will generate a constituent codeword of parity weight approximately proportional to m . Hence, for example, if the interleaver π maps one such input vector \mathbf{u} into another vector $\pi(\mathbf{u})$ of the same type, as in Figure 6 (a), the corresponding turbo codeword will also have a low overall weight. The approach followed in [40] is to show that for any interleaver π , there must exist some short loops of these two types. One important theoretical consequence of this is that the normalized minimum distance d_C/N of turbo code as described in this paper approaches zero for large N , in contrast to the best possible block codes, according to the Varshamov-Gilbert bound [5]. In practice, observed minimum distances d_C of moderate length turbo codes are much lower than Breiling and Huber's upper bounds. Therefore, all turbo codes will display a significant error floor.

III. Iterative Decoding

In Section II, the analysis of the error probability was made under the assumption of maximum likelihood decoding, which in a strict sense unfortunately seems to be infeasible. However, Berrou and his colleagues *re-invented* (see Section V) *iterative decoding*, which for moderate to high SNRs appears to perform very close to maximum likelihood decoding, although a formal proof to this effect has yet to be presented. However, we will return to this issue in Section IV.

Iterative decoding relies on cooperation between the decoding modules of the two constituent codes involved. Each of the decoding modules uses the available information; namely the received values for each transmitted symbol, and certain *a priori* information presented by the other decoder module. To motivate this discussion, let us return briefly to the situation of Figure 1. In this example, *a priori* information was assumed to aid the process of decoding. How-

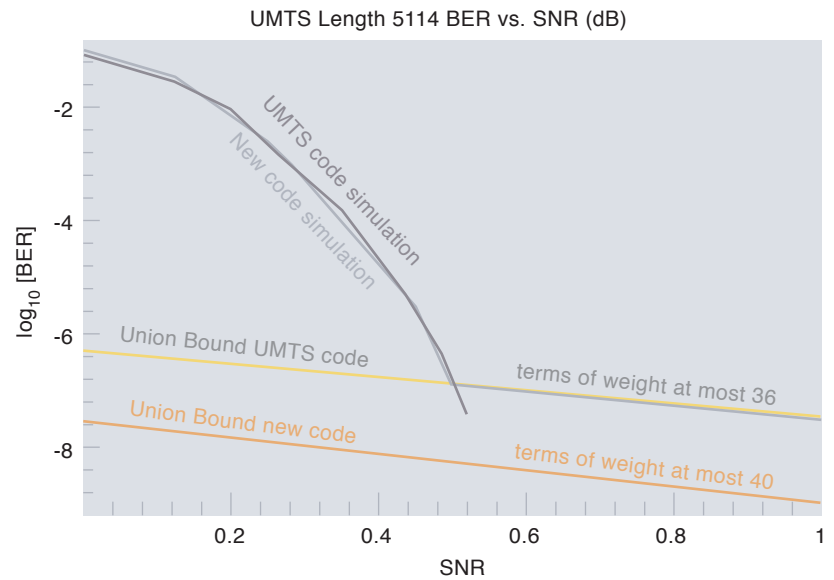


Figure 5 Simulation results and bounds for information length 5114

ever, as the story sadly demonstrates, “incorrect” *a priori* information can end up being disastrous for the decoding process. Of course, in a decoding situation there is no way ahead of time to determine which *a priori* information is “correct” and which is not. Iterative decoding works by using *extrinsic information*, to be defined below, instead of complete information as a connection between the SISO decoders. During the course of many iterations, on average the contributions of correct extrinsic information will outweigh the negative effects of incorrect extrinsic information.

Iterative decoding is presented in Figure 7. Consider a stochastic variable X that takes a value x with some known probability $P(x)$ (perhaps conditioned on some specific event). It will be con-

Figure 6 Low weight errors made up of (a) weight two inputs (b) weight four inputs. Here, $p = 3$

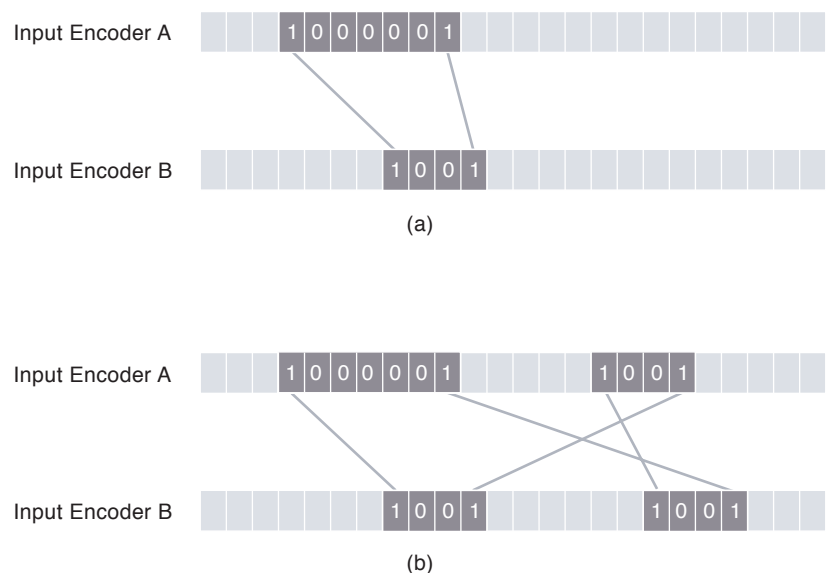
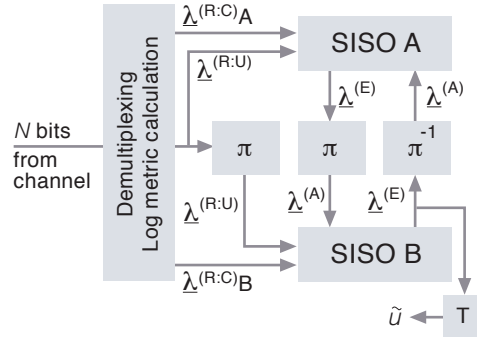


Figure 7 Turbo decoding



venient to represent this probability in terms of a *log-likelihood ratio* (LLR), i.e. a function of the type

$$\lambda(x) = \log \frac{P(x)}{1 - P(x)}. \quad (10)$$

Abusing notation but convening to the literature, we will sometimes refer to LLRs as *information*³⁾ (i.e. channel information, *a priori* information, extrinsic information, and complete information) on the given bit x . Using an LLR instead of a probability means that operations become additive rather than multiplicative.

Let $\underline{\lambda}(x)$ denote a block of LLRs, i.e. $(\lambda(x_1), \dots, \lambda(x_m))$ for some vector (x_1, \dots, x_m) . The receiver de-multiplexes the received sequence into three blocks of LLRs, $\underline{\lambda}^{(R:U)}$, $\underline{\lambda}^{(R:C)A}$, and $\underline{\lambda}^{(R:C)B}$, representing the channel information on the transmitted information symbols and the parity check symbols in the two constituent codes, respectively. The determination of the probability distributions requires an accurate estimate of the received SNR, but it is not particularly difficult to obtain such an estimate when the channel is stable. A punctured and non-transmitted symbol is represented by a zero in the corresponding LLR block.

For each constituent code, a *soft-in-soft-output* (SISO) decoder, to be discussed below, produces the complete *a posteriori* information $\underline{\lambda}^{(\text{Complete})}$ on the block of transmitted symbols, based on the *a priori* information $\underline{\lambda}^{(A)}$, the channel information $\underline{\lambda}^{(R)}$, and the structural relationship between the bits as imposed by the constituent code. On a block level,

$$\underline{\lambda}^{(\text{Complete})} = \underline{\lambda}^{(A)} + \underline{\lambda}^{(R)} + \underline{\lambda}^{(E)}. \quad (11)$$

Subtracting $\underline{\lambda}^{(A)}$ and $\underline{\lambda}^{(R)}$ from $\underline{\lambda}^{(\text{Complete})}$, we obtain the *extrinsic* information $\underline{\lambda}^{(E)}$, which can

be regarded as the amount of information gained by this pass through the decoder. The extrinsic information $\lambda^{(E)}(u_j)$ on a particular information bit u_j is independent of the *a priori* information $\lambda^{(A)}(u_j)$ on that same bit.

At the start of the decoding process, the blocks $\underline{\lambda}^{(R:U)}$ and $\underline{\lambda}^{(R:C)A}$ are submitted to SISO A, while the initial *a priori* information $\underline{\lambda}^{(A)}$ is zero for all symbols. Secondly, the blocks $\underline{\lambda}^{(R:U)}$ (after an appropriate interleaving) and $\underline{\lambda}^{(R:C)B}$ are presented to SISO B, together with the extrinsic information $\underline{\lambda}^{(E)}$ from the previous step, which is now used after the interleaving as a priori information $\underline{\lambda}^{(A)}$ for SISO B. Subsequently, the extrinsic information from SISO B is in turn de-interleaved and presented to SISO A as a priori information for the second round of decoding. As the decoding progresses, the extrinsic information is gradually refined, and information on the whole block is used to produce a final estimate on each user bit u_j , $0 \leq j < K$. Technically, one can argue that a threshold decision on the complete information $\lambda_j^{(\text{Complete})}$ should be used to obtain the estimate \tilde{u}_j on bit u_j . In practice, after some iterations it seems to make little difference if the current extrinsic information $\lambda_j^{(E)}$ is used in place of the complete information.

III.A The SISO Blocks and the BCJR Algorithm

The iterative decoding algorithm relies heavily on the SISO blocks. Several approaches to dealing with the SISO approach have been developed, including the maximum a priori decoding invented by Bahl, Cocke, Jelinek, and Raviv in [4], subsequently known as the *BCJR* algorithm, and *soft-output Viterbi decoding* [15]. Below follows a description of the BCJR algorithm, in the additive version as presented by Benedetto et al. [16]. This version produces the extrinsic information directly.

The BCJR algorithm, like most SISO algorithms, requires a *trellis* [18] representation of the code, see Figure 8. A trellis is a directed graph. The set of nodes or *states* of the graph can be partitioned into subsets $S_0, S_1, \dots, S_{n-1}, S_n$. The states of S_j belong to the j -th *depth*. For a block code, and hence for a terminated convolutional code as the constituent codes of a turbo code, the initial and final depths contain one state each, so $S_0 = \{s_0\}$ and $S_n = \{s_n\}$. An edge \mathbf{e} from a state in S_{i-1} terminates at a state in S_i , $1 \leq i \leq n$. Let $s^S(\mathbf{e})$ and $s^E(\mathbf{e})$ be the states where \mathbf{e} starts and ends, respectively.

³⁾ This informal concept of information should not be confused with the mathematically defined information function of information theory, which is used in the discussion of density evolution in Section IV.

For the *terminated* trellis of a binary rate 1/2 recursive convolutional constituent code, $n = K + v$. The trellis in Figure 8 has $K = 3$ and $v = 2$. During the first K depths there are two edges leaving each state, but starting at the K -th depth there is just one edge leaving each state, giving a unique terminating path leading to the ending state s_n at depth n . Each edge \mathbf{e} is associated with the following labels:

- $u(\mathbf{e})$, the information bit associated with \mathbf{e} . This is shown by the color in Figure 8 (grey is zero, orange is one).
- $c(\mathbf{e})$, the parity check symbol associated with \mathbf{e} . This is not shown in Figure 8.

Consider a path starting in s_0 and terminating in s_n in the constituent code trellis. Collecting the pair of edge labels at each edge, we obtain an *edge label sequence*, which corresponds to a codeword of the constituent code. The constituent code consists of the set of 2^K edge label sequences obtained by traversing all possible paths from s_0 to s_n .

The algorithm requires two passes (or recursions), one forward pass through the trellis where certain values regarding the j -th depth of the trellis are computed in a recursion based on the same values at the $(j - 1)$ -th depth; and one backward pass. In the forward recursion, the decoder starts the recursion with $\alpha_0(s_0) = 0$ and computes, for depth $j = 1, \dots, n$ and for each state s at depth j , the value

$$\alpha_j(s) = \max_{\mathbf{e}: s^S(\mathbf{e})=s} \left\{ \underbrace{\alpha_{j-1}(s^S(\mathbf{e}))}_{\text{preceding state}} + \underbrace{u(\mathbf{e})\lambda_j^{(A)}}_{\text{a priori value}} + \underbrace{u(\mathbf{e})\lambda_j^{(R:U)} + c(\mathbf{e})\lambda_j^{(R:C)}}_{\text{channel information}} \right\} \quad (12)$$

where $\max^* a_j$ operating on T numbers a_1, \dots, a_T is the *sum*-operation

$$\log \left[\sum_{t=1}^T e^{a_t} \right] \quad (13)$$

The quantity $\alpha_j(s)$ is an LLR representing the probability that the constituent encoder is in trellis state s at time j , conditioned on the available preceding channel information and *a priori* information pertaining to all symbols until time j . To avoid numerical problems, the values $\alpha_j(s)$ are normalized at each depth j by finding $m = \max_s \alpha_j(s)$, and subtracting m from $\alpha_j(s)$ for all states s at depth j .

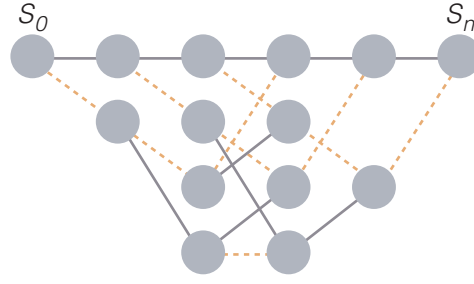


Figure 8 A trellis

The backward pass similarly starts the recursion with $\beta_n(s_n) = 0$ and calculates

$$\beta_j(s) = \max_{\mathbf{e}: s^S(\mathbf{e})=s} \left\{ \beta_{j+1}(s^E(\mathbf{e})) + u(\mathbf{e})\lambda_{j+1}^{(A)} + u(\mathbf{e})\lambda_{j+1}^{(R:U)} + c(\mathbf{e})\lambda_{j+1}^{(R:C)} \right\} \quad (14)$$

for all states s at depth j , $j = n - 1, \dots, 1$. The quantity $\beta_j(s)$ is an LLR representing the probability that the constituent encoder is in trellis state s at time j , conditioned on the available channel information and *a priori* information on all symbols after time j . The values $\beta_j(s)$ are normalized in the same way as $\alpha_j(s)$.

The final step of the SISO module is to compute, for all j , $0 \leq j < K$, the extrinsic log metric of the j -th information bit, $\lambda_j^{(E)}$. For $u = 0, 1$, compute

$$\mu(u, j) = \max_{\mathbf{e}: u(\mathbf{e})=u} \left\{ \alpha_{j-1}(s^S(\mathbf{e})) + \beta_j(s^E(\mathbf{e})) + u(\mathbf{e})\lambda_j^{(R:U)} + c(\mathbf{e})\lambda_j^{(R:C)} \right\}. \quad (15)$$

Then

$$\lambda_j^{(E)} = \mu(1, j) - \mu(0, j). \quad (16)$$

Note that the extrinsic information $\lambda_j^{(E)}$ on the j -th information bit is independent of the *a priori* value $\lambda_j^{(A)}$ on the j -th bit.

III.B Termination Criteria

Experience, backed by the density evolution arguments to be described in Section IV, shows that the performance of iterative decoding usually improves as the number of iterations increases, up to a certain point. The number of iterations before this convergence is observed depends on the SNR.

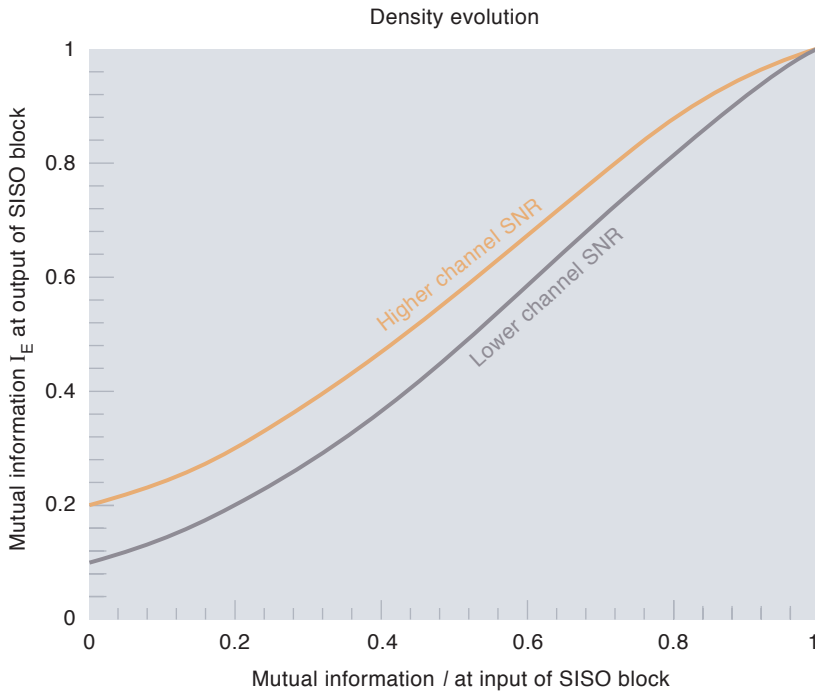


Figure 9 The connection between a priori information and extrinsic information

How many iterations must be carried out? Hagenauer et al. [15] suggested to use a threshold on the *cross entropy* between the extrinsic information produced by the two decoders. Wu et al. [30] suggested to use a threshold on the number of sign differences between these two extrinsic values. To avoid some rare cases where the decoder does not converge, these termination criteria must be accompanied by a maximum limit on the number of iterations. The simulated BER performance of these techniques is close to what is obtained with a virtually unlimited number of iterations. The actual number of iterations depends on the constituent codes as well as on the channel SNR, as can also be deduced from the density evolution arguments in the next section. For “typical” constituent codes, [30] claims that the average number of iterations is small – less than four or five – at SNRs lower than the error floor threshold.

IV. Selection of System Components

The turbo code is made up from the constituent codes, the interleaver, and the puncturing schemes. A complete optimization of puncturing schemes seems to be infeasible; on the other hand, among simpler schemes there does not seem to be much of a difference. We will therefore not discuss the issue of puncturing, and focus on the constituent codes and the interleavers.

IV.A Constituent Codes

As for convolutional codes, there is a trade-off between decoding complexity and performance. Decoding complexity is directly related to trellis complexity [18]. At the error floor, complex constituent trellises generally offer better performance, but in the waterfall region the picture is not so clear.

IV.A.1 Performance at the Error Floor

In [20], a computer search is used to find good convolutional constituent codes. The goal of this search was to find recursive convolutional encoders which offer high weight codewords for all weight two input vectors, due to the lessons learnt from equation (9). However, for non-random interleavers, weight two input vectors may be less important. Another design approach is to use codes whose codeword weights grow fast with the *active* length of the input vectors, i.e. the number of successive time instants when the encoder is not in the zero state. This corresponds to maximizing the minimum average cycle weights of the constituent codes.

IV.A.2 Performance in the Waterfall Region

Recently [33], [35], [39], *density evolution* considerations have been introduced as a tool for studying the iterative decoding procedure. In this approach, it is assumed that each SISO produces extrinsic information which obeys a Gaussian probability density function, and which is not correlated with the observed received values. These assumptions are reasonable for large interleavers [39].

The technique can be applied to many channel models. We assume an AWGN channel. For a fixed channel SNR, define $I_A = I(\underline{\lambda}^{(A)}, \mathbf{u})$ to be the mutual information between the *a priori* values and the transmitted information.⁴⁾ Similarly, we can define $I_E = I(\underline{\lambda}^{(E)}, \mathbf{u})$ to be the mutual information between the extrinsic values produced by the SISO and the transmitted information. Now, consider the function $I_E = F(I_A, SNR)$. Analytical expressions for this function have been obtained for simple constituent codes [33]. For a general code, analytical expressions are unknown, but the relationships can easily be obtained by computer simulation. This computer simulation is much simpler [39] than a turbo code BER simulation, which requires the observation of a large number of very rare error events. In Figure 9 a typical function $I_E = F(I_A, SNR)$ is shown for two different SNRs.

⁴⁾ At this point we will not go into the finer details of information theory. For a presentation of information theory, see any textbook, for example [9]. Note that a mutual information value of 1 means that knowledge of one parameter exactly determines the other. A mutual information value of 0 means that the two parameters are statistically independent.

Figure 10 contains an EXIT (extrinsic information transfer) chart. The functions F for the first SISO and F^{-1} for the second SISO, for some fixed SNR, are plotted in the same coordinate system. The iterative decoding procedure follows a trajectory between these two curves. As shown in Figure 10, at low SNR, the two curves intersect so that decoding beyond a few iterations will not lead to improved decoding results. At moderate SNR, there is an open tunnel between the two curves, allowing decoding to proceed. The smallest SNR for which the tunnel is open is called the *pinch-off limit* in [39], which corresponds to the SNR values at which the waterfall region starts. The EXIT charts can be used for multiple purposes, such as searching for and evaluating good constituent codes, and predicting (reasonably accurately) the BER after a prescribed number of decoding iterations. This technique also presents a strong argument that turbo coding is very close to optimal at high SNR.

Note that the density evolution technique does not take into consideration the effect of a limited minimum distance. Thus it cannot be applied to determine the error floor.

IV.B Interleavers

In principle, any permutation of the K input bits can serve as an interleaver, but all permutations are not equally good. Before we proceed to performance related issues, note that any permutation can be represented by a $K \times K$ lookup table. For large K , or in the case where K is adaptive to channel conditions or can be selected by the user (within a range), lookup tables may be inconvenient. In such cases it would be nice to have a simple and fast deterministic algorithm to specify the interleaver. For example, a simple block interleaver provides maximum spreading of the interleaved symbols, which according to both the following subsections is an advantage. However, the regularity of this construction also seems to be disadvantageous both in the waterfall region and especially at the error floor region. As of this writing, the known simple deterministic algorithms produce interleavers with a worse performance than the interleavers discussed below.

The purpose of the interleaver is basically to optimize the performance of the turbo codes. We will consider the waterfall region and the error floor region separately.

IV.B.1 Performance in the Waterfall Region

The density evolution arguments rely on the information that is passed between the constituent SISO modules to be independent. Since these SISO modules are linked through the inter-

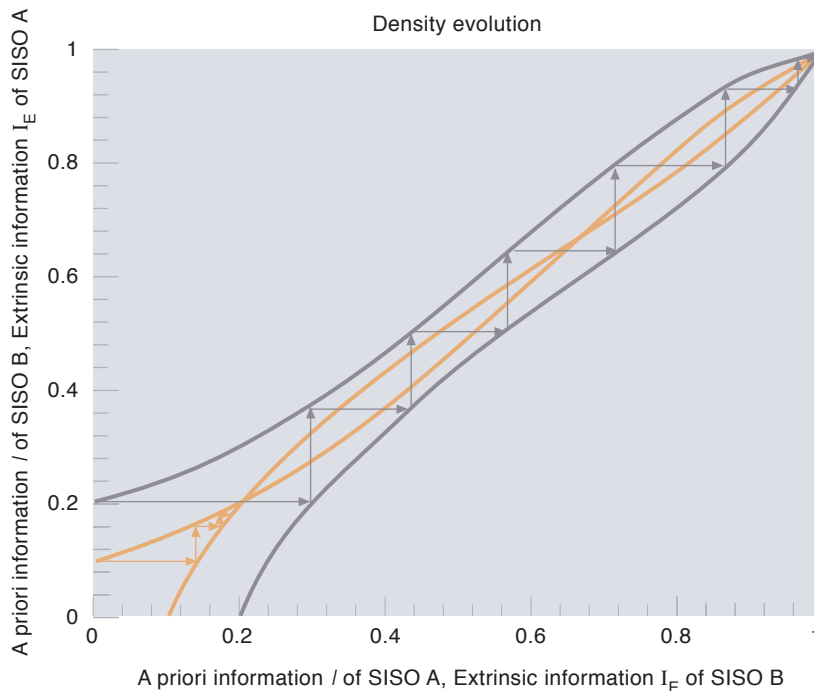


Figure 10 Exit chart

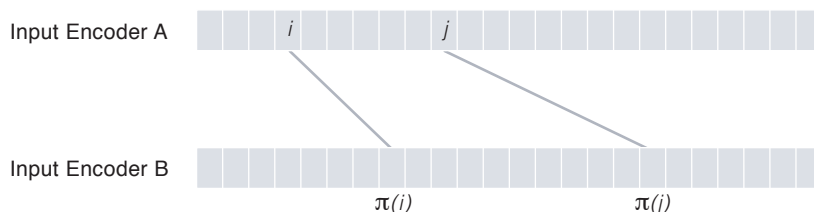
leavers, this assumption is not entirely correct. A *first order cycle* is made up of any two numbers $i, j, 0 \leq i < j < K$. The length of the cycle is defined as $l = l(i, j) = j - i + |\pi(i) - \pi(j)|$. One such cycle is shown in Figure 11. If l is small, the independence assumption is jeopardized. Indeed, the negative impact on the waterfall performance in an interleaver with many short cycles can be shown by density evolution arguments as well as by simulation.

Hokfeldt et al. [26] designed interleavers with the aim of minimizing the correlation between the extrinsic information produced by successive SISO iterations. The method seeks to minimize the number of short first order cycles. There does not seem to be a significant difference in the waterfall region between these interleavers and those designed in [42], designed next.

IV.B.2 Performance in the Error Floor Region

In the error floor region, the key parameters determining performance are the code's minimum distance, and the number of codewords of minimum distance. The interleaver can be

Figure 11 Short interleaver cycles



designed with the aim to optimize these parameters.

Interestingly, the simplest step is again to make sure that there are no short first order cycles in the interleaver. Dolinar and Divsalar suggested to use s -random interleavers [13]. Their s -random interleavers do not contain interleaver mappings $\pi(i), \pi(j)$ if

$$|i - j| < s \text{ and } |\pi(i) - \pi(j)| < s, \text{ simultaneously.} \quad (17)$$

These are pseudo-randomly generated interleavers, where interleaver mappings $\pi(i)$ are added one by one and discarded during the generation process if they violate (17) together with any of the existing mappings $\pi(j)$. Crozier [42] redefined the s -random property slightly, by requiring that $l = |i - j| + |\pi(i) - \pi(j)| \geq s$, and presented two efficient ways to generate powerful interleavers.

Andrews et al. [21] suggested to also remove short cycles of the type shown in Figure 6 (a). Breiling et al. [32] designed interleavers for specific pairs of constituent codes, by explicitly avoiding loops that involved known low weight codewords in both constituent codes. The “other” interleaver in Figure 5 was found by using these techniques in combination with the ones in [42].

V. Variations, Generalizations, Applications and Limitations

V.A Variations

Codes can be concatenated in other ways than by parallel concatenation, and still be decoded by the turbo decoding method. This can include two or more constituent codes. Serial concatenation is one possibility.

V.A.1 Serial Concatenation

Serial concatenation [19] is shown in Figure 12. Such constructions can also be decoded with an iterative decoding method. For details see [19]. Serially concatenated codes tend to have larger minimum distances than parallel concatenated codes of comparable complexity and rate. However, research results so far suggest that the turbo decoding algorithm also seems to be less efficient in this case.

V.A.2 Repeat-Accumulate Codes

Divsalar et al. [23] suggested a particularly simple serial concatenation: The inner “code” is a rate 1 accumulator mapping, i.e. its single output bit at time j is simply the modulo 2 sum of all input bits so far. The accumulator mapping can be described by a 2-state trellis, to which a SISO module is applied. The outer code is a rate 1/3 repetition code: each information bit is copied twice in the output. This simple construction performs surprisingly well, especially in the waterfall region. This can be explained by density evolution arguments [33].

V.B Generalizations

A turbo code can be viewed as a low-density parity-check (LDPC) code, introduced already in 1962 by Gallager [3], and reinvented by Tanner in 1981 [6]. Apparently, technology was not ripe for these ideas at the time when they were first introduced.

Wiberg et al. [12] observed that turbo codes, as well as LDPC codes, can be described by a *Tanner graph*. Turbo decoding, as well as iterative decoding of LDPC codes, proceeds as a message passing algorithm on this graph. This graph and the decoding algorithm acts by *factoring* the decoding problem. Generalizing these concepts, we arrive at *factor graphs*, which cover a vast number of problems, from many application areas, and their solutions, as special cases. See [34] for more details.

In the message passing algorithms, optimal *scheduling* of the messages is a problem yet to be solved. This has led to the introduction of parallel message passing algorithm, and to the extreme cases of parallelism: Analog decoders [22], [25].

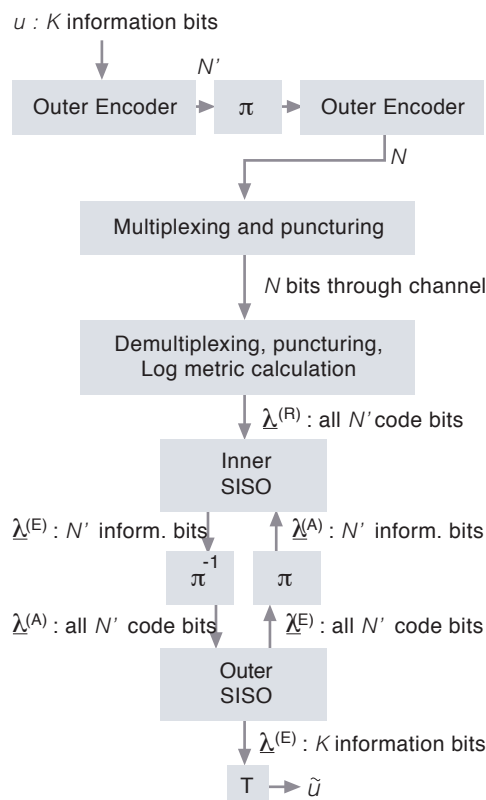


Figure 12 Serial concatenation

V.C Applications

This paper has focused on turbo codes applied to AWGN channels with BPSK modulation. Turbo codes can of course be applied to a variety of coding and communication situations. We review some of these below.

- *Coded modulation* is possible with turbo codes as well. In contrast to traditional methods, like trellis coded modulation [7], set partitioning cannot be readily applied. Some challenges remain, concerning issues such as how to map the binary encoded bits to signal constellation points, or rotation invariance. Still, turbo coded modulation produce good results [29].
- Turbo codes have been applied with success to *fading channels* [29]. This usually requires a combined iterative detection/decoding scheme where equalization [25] and/or estimation of channel parameters [38] are included in the decoding process.
- An example of a completely different application is the use of turbo decoding in correlation attacks on cryptographic functions [27], [28].

McEliece [41] suggests that turbo-like codes will work well in most communication situations.

V.D Limitations

Turbo coding performs close to the Shannon bound even when a very small BER is required. The relatively small minimum distance presents some problems, though. This means that to achieve a small BER at low SNR, a certain minimum information length N is required. Since turbo decoding processes one block at a time, this can cause a decoding delay beyond what is tolerable for some applications.

References

- 1 Asbjørnsen, P C, Moe, J. God dag mann! – Økseskaft! In: *Norske Folkeeventyr*, 1841 – 1844. (In Norwegian)
- 2 Shannon, C. A Mathematical Theory of Communication. *Bell System Tech. J.*, 27, July and October, 379–423, 623–656, 1948.
- 3 Gallager, R. Low-density parity-check codes. *IRE Transactions on Information Theory*, IT-8 (1), 21–28, 1962.
- 4 Bahl, L R et al. Optimal decoding of linear codes for minimizing symbol error rate. *IEEE Transactions on Information Theory*, IT-20 (2), 284–287, 1974.
- 5 MacWilliams, F J, Sloane, N J A. *The theory of error-correcting codes*. Amsterdam, North-Holland, 1977.
- 6 Tanner, M. A recursive approach to low complexity codes. *IEEE Transactions on Information Theory*, IT-27 (5), 533–547, 1981.
- 7 Ungerboeck, G. Channel coding with multi-level/phase signalling. *IEEE Transactions on Information Theory*, IT-28 (1), 55–67, 1982.
- 8 Lin, S, Costello, D. *Error Control Codes*. Englewood Cliffs, Prentice Hall, 1983.
- 9 Johannesson, R. *Informationsteori : Grundvalen för (tele-)kommunikation*. Lund, Studentlitteratur, 1988. (In Swedish).
- 10 Blahut, R. *Digital Transmission of Information*. New York, Addison-Wesley, 1990.
- 11 Berrou, C, Glavieaux, A, Thitimajshima, P. Near Shannon limit error correcting coding and decoding : Turbo-codes. In: *Proc. ICC'93*, Geneva, Switzerland, May 1993, 1064–1070.
- 12 Wiberg, N, Loeliger, H-A, Koetter, R. Codes and iterative decoding on general graphs. *European Transactions on Telecommunications*, 6 (5), 513–525, 1995).
- 13 Dolinar, S, Divsalar, D. *Weight distributions for turbo codes using random and nonrandom permutations*. Pasadena, CA, USA, Jet Propulsion Lab (JPL), 1995. (TDA Progress report, 42-122.)
- 14 Benedetto, S, Montorsi, G. Unveiling turbo codes: some results on parallel concatenated coding schemes. *IEEE Transactions on Information Theory*, 42 (3), 409–428, 1996.
- 15 Hagenauer, J, Offer, E, Papke, L. Iterative decoding of binary block and convolutional codes. *IEEE Transactions on Information Theory*, 42 (3), 429–445, 1996.
- 16 Benedetto, S et al. A soft-input soft-output maximum a posteriori (MAP) module to decode parallel and serial concatenated codes. Pasadena, CA, USA, Jet Propulsion Lab (JPL), 1996. (TDA Progress report, 42-127.)
- 17 Pless, V, Huffman, W C (eds.). *Handbook of Coding Theory*. Amsterdam, North-Holland, 1998.

- 18 Vardy, A. Trellis structure of codes. In: Pless, V, Huffman, W C (eds.). *Handbook of Coding Theory*. Amsterdam, North-Holland, 1998. 1989–2118.
- 19 Benedetto, S et al. Serial Concatenation of Interleaved Codes : Performance Analysis, Design, and Iterative Coding. *IEEE Transactions on Information Theory*, 44 (3), 909–926, 1998.
- 20 Benedetto, S, Garelo, R, Montorsi, G. A search for good convolutional codes to be used in the construction of turbo codes. *IEEE Transactions on Communications*, 44 (9), 1101–1105, 1998.
- 21 Andrews, K S, Heegard, C, Kozen, D. Interleaver design methods for turbo codes. In: *Proc. IEEE Int. Symposium on Information Theory*, Cambridge, MA, USA, 1998, 420.
- 22 Loeliger, H-A et al. Iterative sum-product decoding with analog VLSI. In: *Proc. IEEE Int. Symposium on Information Theory*, Cambridge, MA, USA, 1998, 146.
- 23 Divsalar, D, Jin, H, McEliece, R J. Coding Theorems for ‘turbo-like’ codes. In: *Proc. 1998 Allerton Conference on Communications*, Allerton, IL, USA, Sept. 1998.
- 24 Heegard, C, Wicker, A B. *Turbo Coding*. Boston, Kluwer, 1999.
- 25 Hagenauer, J et al. Decoding and equalization with analog non-linear networks. *European Transactions on Telecommunications*, 10 (5), 659–690, 1999.
- 26 Hokfeldt, J, Edfors, O, Maseng, T. Interleaver design for turbo codes based on the performance of iterative decoding. In: *Proc. IEEE International Conference on Communications*, Vancouver, BC, Canada, June 1999.
- 27 Johansson, T, Jönsson, F. Fast correlation attacks based on Turbo code techniques. In: *Proceedings of Crypto’99*, Santa Barbara, CA, USA, August 1999. Lecture Notes in Computer Science, Berlin, Springer, 1999.
- 28 Fossorier, M P C, Mihaljevic, M J, Imai, H. Critical noise for convergence of iterative probabilistic decoding with belief propagation in cryptographic applications. In: *Proceedings of AAECC’13*, Honolulu, HI, USA, November 1999. Lecture Notes in Computer Science, Berlin, Springer, 1999.
- 29 Vucetic, B, Yuan, J. *Turbo Codes : Principles and Applications*. Boston, Kluwer, 2000.
- 30 Wu, Y, Woerner, B D, Ebel, W J. A simple stopping criterion for turbo decoding. *IEEE Communication Letters*, 4 (8), 258–260, 2000.
- 31 Breiling, M, Huber, J B. A method for determining the distance profile of turbo codes. In: *Proc. of 3rd ITG conference on source and channel coding*, Munich, Germany, January 2000.
- 32 Breiling, M, Peeters, S, Huber, J B. Interleaver design using backtracking and spreading methods. In: *Proc. IEEE Int. Symposium on Information Theory*, Sorrento, Italy, June 2000, 451.
- 33 Divsalar, D, Dolinar, S, Pollara, F. Low complexity turbo-like codes. In: *Proc. 2nd Int. Symposium on Turbo Codes*, Brest, France, September 2000, 73–80 .
- 34 Kschischang, F R, Frey, B J, Loeliger, H-A. Factor graphs and the sum-product algorithm. *IEEE Transactions on Information Theory*, 47 (2), 498–519, 2001.
- 35 Richardson, T, Urbanke, R. The capacity of low-density parity-check codes under message-passing decoding. *IEEE Transactions on Information Theory*, 47 (2), 599–618, 2001.
- 36 Chung, S Y et al. On the design of low-density parity-check codes within 0.0045 dB of the Shannon limit. *IEEE Communications letters*, 5 (2), 58–60, 2001.
- 37 Garelo, R G, Pierleoni, P, Benedetto, S. Computing the free distance of turbo codes and serially concatenated codes with interleavers: Algorithms and applications. *IEEE Journal of Selected Areas in Communications*, 19 (5), 800–812, 2001.
- 38 Valenti, M C, Woerner, B D. Iterative channel estimation and decoding of pilot symbol assisted turbo codes over flat-fading channels. *IEEE Journal on Selected Areas in Commun.*, 19 (9), 1697–1705, 2001.
- 39 ten Brink, S. Convergence behavior of iteratively decodes parallel concatenated codes. *IEEE Transactions on Communications*, 49 (10), 1727–1737, 2001.

- 40 Breiling, M, Huber, J B. Combinatorial analysis of the minimum distance of turbo codes. *IEEE Transactions on Information Theory*, 47 (7), 2737–2750, 2001.
- 41 McEliece, R J. Are turbo-like codes effective on non-standard channels? *IEEE Information Theory Newsletter*, 51 (4), 2001. (Presented at IEEE Int. Symposium on Information Theory, Washington, D.C., USA, June, 2001.)
- 42 Crozier, S N. *New High-Spread High-Distance Interleavers for Turbo-Codes*. Preprint, 2001.
- 43 Rosnes, E, Ytrehus, Ø. *Algorithms for turbo code weight distribution calculation with applications to UMTS codes*. To be presented at IEEE Int. Symposium on Information Theory, Lausanne, Switzerland, July 2002.

Theory and Practice of Error Control Coding for Satellite and Fixed Radio Systems

PÅL ORTEN AND BJARNE RISLØW



Pål Orten (35) is Research Manager at Nera Research. He received his *Siv.Ing.* degree from the Norwegian University of Science and Technology in 1989, and his PhD from Chalmers University of Technology in 1999. From 1990 to 1995 he was Research Scientist at ABB Corporate Research and from 1995 at Nera Research (interrupted by PhD studies). In addition to various research and development activities in channel coding and signal processing at ABB and Nera, he also participated in the European research project FRAMES, which resulted in the definition of the 3G UMTS system in ETSI. His current research has focus at fixed wireless communications and mobile satellite communications.

p.al.orten@research.nera.no



Bjarne Risløw (38) is a Research Scientist at Nera Research. He received his MSc in Electrical Engineering at the Norwegian Institute of Technology in 1988. He has worked as a Research Scientist at SINTEF DELAB (1989–1992), ABB Corporate Research (1993–1994) and from 1995 at Nera Research. At ABB and Nera he has worked with the development of terminals and earth stations for the Inmarsat systems. In the last years he has been mostly involved in wireless broadband access project, DVB-RCS and LMDS. His main field of interest is with terminal technologies and in particular modem technology, coding and synchronization.

bjarne.risløw@research.nera.no

Strong “state-of-the-art” Forward Error Correction (FEC) coding has been extensively applied for both satellite and fixed radio communication. With the discovery of Turbo codes, performance close to the capacity limits for moderate bit error rates became possible. In this paper we discuss various aspects of the capacity theorem and the capacity limits. We describe and present results for some error control coding schemes that are currently applied in satellite and fixed radio systems, and compare these results with the theoretical limits. Finally, we describe and evaluate promising new coding schemes that have not yet been applied in commercial systems.

1 Introduction

By the widespread use of mobile cellular phones, radio communications have become a natural part of life for many people around the world. This has been made possible by advanced technical research enabling cost-effective implementation and reliable communication in a rather harsh environment. In general a radio channel is exposed to many degrading and disturbing effects like interference, multipath propagation causing fading, Doppler shifts, and thermal noise. Obviously, to be able to operate a communication system under such conditions, some method or scheme for error control is required. Many of these advanced error control schemes were developed for deep space communications and satellite communications where power (and to some extent bandwidth) limits the performance. Other advances in coding theory have been made for applications where bandwidth efficiency is very important. Examples of such systems are cable modems and radio relay communications.

An example of a satellite communication link is shown in Figure 1. Satellite communication systems

are characterised by an extremely large distance between transmitter and receiver, since the communication signal is transmitted via a satellite located up to 36,000 km (for Geosynchronous Earth Orbits – GEO) above the earth surface. Since the intensity of electromagnetic waves (in free space) decays with the square of the radio path length, such a system is clearly power limited. Therefore, satellite systems normally need line of sight to have sufficient link margin for reliable operation. Naturally, sophisticated error control coding schemes are necessary for acceptable performance, as well as for limiting the power consumption in both satellite and terminal. This is especially important for mobile terminals. Another major problem with the long signal path of GEO satellites is the corresponding long delay. This delay might reduce the perceived quality of speech services and cause problems for delay sensitive data communication services. Thus, satellite systems have also been designed with satellites in lower earth orbits. For such systems, the delay is reduced, but other problems like more Doppler and need for hand-over between satellites occur. A common and accepted channel model for satellite

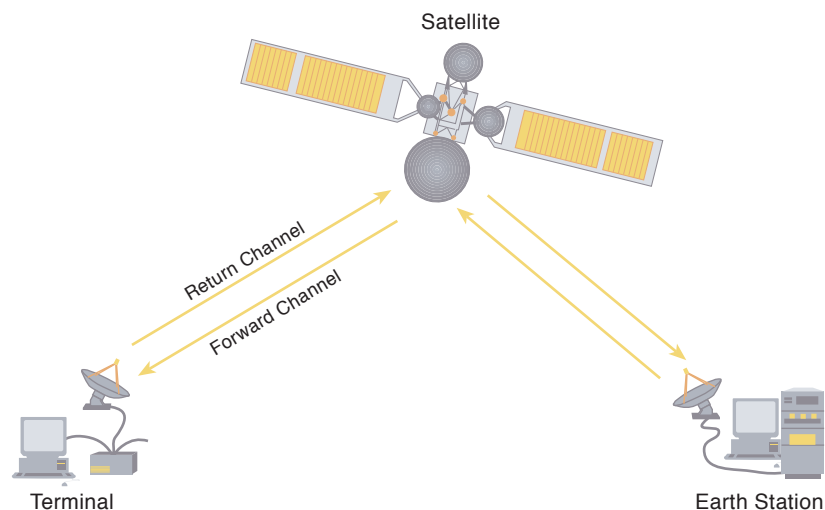


Figure 1 Satellite System

systems is the Ricean channel model. With handheld phones, and thus rather non-directive antennas, the scattered components may be significant. However, with rather directive antennas, and slowly moving or fixed terminals, the direct ray is very strong compared to the diffuse scattered components. The channel is then quite close to an Additive White Gaussian Noise (AWGN) Channel. For maritime or aeronautical applications these assumptions may not hold, and other fading models must be used. More information about satellite channels can be found for instance in [1].

Figure 2 shows a radio link communication system with some high capacity radio hops (typically STM-1 data rates¹⁾ bringing the communication to a city where the traffic is further distributed to the user by cable or radio transmission. Radio link communication systems are rather different from satellite communication systems, since they are typically used instead of fibre cables in areas where it is impractical or not cost effective to deploy fibre cables. Radio relay systems therefore have to be highly reliable, provide rather high data rates, and also a very low bit error rate. The goal for the radio system design is then to provide close to fibre quality and speed. Obviously, this sets strong requirements on the channel coding that may be applied. Radio links operate with line of sight communications, with antennas normally mounted several meters above the ground. Nevertheless, reflections may occur from oceans and/or mountains resulting in a multipath fading channel. Also precipitation may cause the signal to fade. A much used channel model for radio link communication is a two-ray model, where the rays are separated by 6.3 ns.

In this paper we focus on coding for satellite communications and radio link communication systems. Initially, we present the capacity limits of Shannon, as well as the practical limits for a given code rate, modulation method and block length. We then study coding schemes that have been applied in many satellite and radio link systems, and show examples of their performance. We then proceed with a presentation of Turbo codes that are now implemented in several recently designed satellite systems. Next, we look at coding schemes and performance for high spectral efficiency radio systems. Furthermore, we present coding schemes that have not yet been applied in commercial systems, including two block code based coding schemes with iterative decoding, Turbo Product Codes and Low Density parity Check Codes. We believe these are strong candidates to be implemented

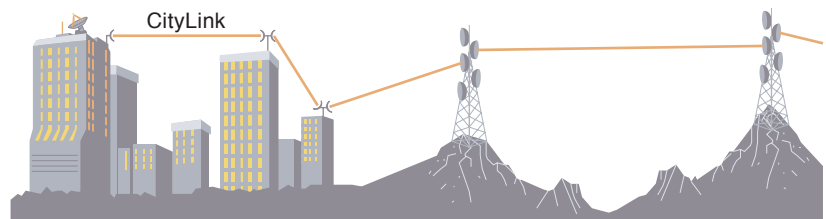


Figure 2 Radio link communication system

in future satellite and radio systems. Finally, the paper presents a comparison of some of the coding schemes along with a discussion of their properties.

2 The Capacity Theorem and Capacity Limits

2.1 Introduction

Claude Shannon developed the formulas for “the maximum rate of transmission of binary digits over a system when the signal is perturbed by various types of noise” [2]. For the AWGN channel with bandwidth, W , and signal-to-noise ratio, S/N , Shannon showed that it was possible to transmit binary digits at a rate

$$C = W \log_2 \left(1 + \frac{S}{N} \right) \quad (1)$$

with as small probability of errors as desired. This maximum rate is also called the channel capacity. Shannon used a geometrical approach to prove this. Assume that m bits are encoded with $M = 2^m$ different signal functions inside the code sphere $\sqrt{2TWP}$ of radius. P is the average power of the codes. Then, by letting T (code length) go to infinity one would reach the capacity limit. A remarkable point is that the bound holds on average even if the M signals are picked at random within the code sphere.

Many researchers have tried to design random-like codes, but a decoder for a pure random code would be very complex since the decoder would have to compare the received signal against all M possible transmitted sequences. This decoder would generally have to calculate Euclidean distances (use soft decision). With the discovery of Turbo codes one suddenly found a practical way to decode long random-like codes using soft decisions, but still with limited decoding complexity.

Let us take a closer look at the capacity formula. When signalling at a certain rate R_s the signal to noise ratio S/N relates to the energy to noise density for each bit E_b / N_0 as:

¹⁾ STM-1 (Synchronous Transport Module level 1) is the basic transmission rate (155.52 Mbit/s) in SDH (Synchronous Digital Hierarchy).

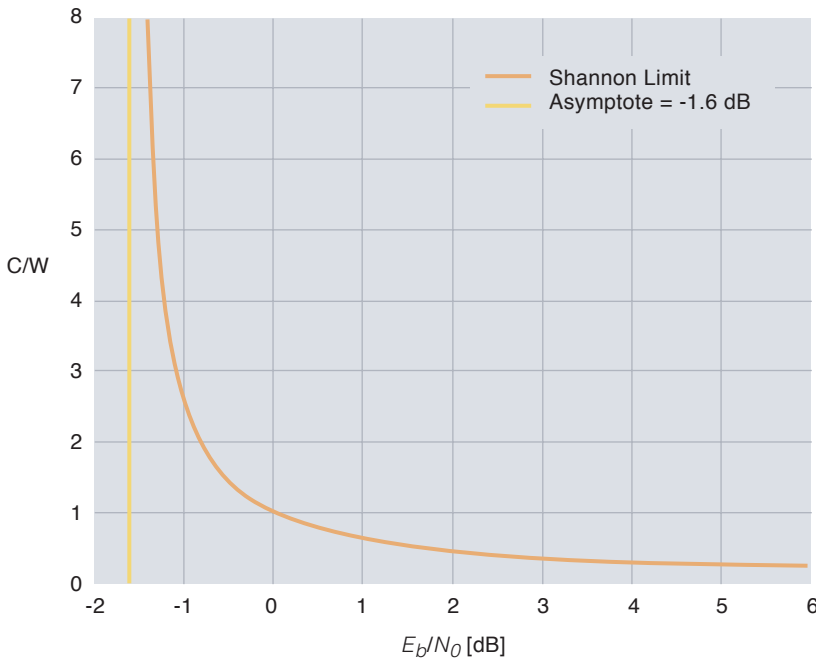


Figure 3 Relationship between the minimum required E_b / N_0 and bandwidth (W/C)

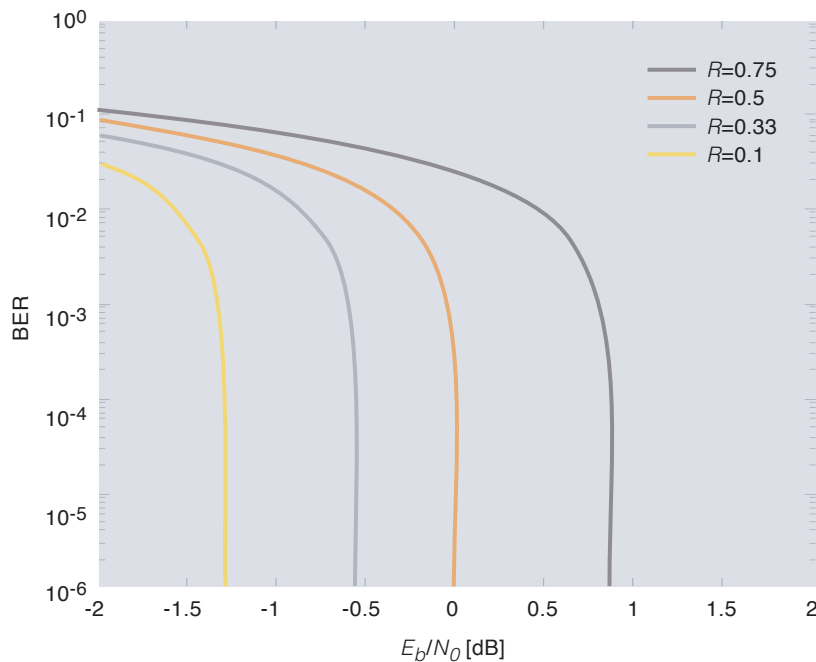
$$\frac{S}{N} = \frac{E_b R_s}{N_0 W} \quad (2)$$

The minimum required E_b / N_0 when transmitting at capacity ($R_s = C$) is then:

$$\frac{E_b}{N_0} = \frac{W}{C} (2^{C/W} - 1) \quad (3)$$

The relationship between the minimum required E_b / N_0 and the relative bandwidth (W/C) is shown in Figure 3.

Figure 4 Minimum bit error rate as function of code rate



It is then easy to see that when the channel bandwidth W goes to infinity the minimum required E_b / N_0 will become $\ln(2)$ or -1.6 dB. This result is not very useful since unlimited bandwidth will not be available. Nevertheless, it is a fundamental limit. Now if we let the $\eta = C/W$ be the spectral efficiency (bit/s/Hz), then Eq. (3) becomes $E_b / N_0 = (2^\eta - 1) / \eta$, which is the unconstrained limit as plotted in Figure 5.

Until a few years ago there was no practical way of getting really close to the Shannon limit. With a certain coding scheme a reasonably good coding gain compared to uncoded modulation could be achieved, but we were still far away from the Shannon limit. As an example, concatenation of an inner convolutional and an outer Reed-Solomon code was still about 2–3 dB away from the capacity limit. However, with the discovery of Turbo codes and iterative decoding one closed the gap and achieved performance within 1 dB of the Shannon limit with practical code lengths.

2.2 Capacity Limit as Function of Code Rate

Eq. (1) can be re-written to take into account the effect of the code rate. If we let $S/N = R * E_b / N_0$ where R is the code rate of the code, i.e. the ratio between the number of information bits and actual number of bits transmitted on the channel, and $W = 1/2T = R_s / 2$ we get the capacity in bits per dimension as:

$$C = \frac{1}{2} \log_2 \left(1 + \frac{2RE_b}{N_0} \right) \quad (4)$$

From the converse to the coding theorem one can derive the inequality, [3]

$$R \cdot (1 - H_b(p)) \leq C \quad (5)$$

where

$$H_b(p) = -p \cdot \ln(p) - [1 - p] \cdot \ln[1 - p]. \quad (6)$$

$H_b(p)$ is the binary entropy function where p is the coded bit error probability. One can now calculate the bit error rate for a specific code rate and E_b / N_0 and the results are shown in Figure 4.

Furthermore, the Shannon limit for the different code rates can be written as:

$$\frac{E_b}{N_0} = \frac{2^{2R} - 1}{2R} \quad (7)$$

From Eq. (8) and Figure 4 we can see that the capacity limit by using 1/2-rate coding is 0 dB.

2.3 Capacity Limit with Constrained Input

The bounds calculated earlier in this document are based on unconstrained input, but most systems use some kind of constrained input like for instance BPSK, QPSK, 8-PSK, or 16-QAM. The channel capacity for equiprobable binary constrained input (± 1), [4], can be written as:

$$C = \frac{1}{2} \int_{-\infty}^{\infty} p(y|1) \log_2 \frac{p(y|1)}{p(y)} dy + \frac{1}{2} \int_{-\infty}^{\infty} p(y|-1) \log_2 \frac{p(y|-1)}{p(y)} dy \quad (8)$$

The probability density function of the received signal including noise, $p(y | \pm 1)$, has mean ± 1 and variance $\sigma^2 = 1 / (2 * E_s / N_0)$. Adjusting for the code rate, we have $E_s / N_0 = R * E_b / N_0$. We can now determine the capacity in bit/s/Hz from Eq. (8). This expression can easily be extended to any square QAM constellation, and the capacity limit for some of these constellations is shown in Figure 5.

If we assume QPSK and a code rate of 1/2 this gives a spectral efficiency of 1 bit/s/Hz. The capacity limit in this case is 0.2 dB while the unconstrained limit from Figure 4 was 0 dB. We see that for higher order modulation schemes the equiprobable square QAM constellation will never reach the Shannon limit, and asymptotically this distance will be 1.53 dB. There are two ways to improve the performance; either by making the probability distribution more Gaussian or by increasing the dimension of the signal constellation. This is also called shaping gain, and at high SNR this gain is separable from the coding gain. Methods exist that can easily obtain 1 dB of the theoretical gain.

2.4 Capacity Limit as Function of Block Length

Finally the actual code length used will limit the capacity. This can be found from the sphere packing bound and the detailed formulas can be found in [5]. In Figure 6 we have plotted the capacity limit as function of block length for 1/2 and 3/4 rate.

Note that the influence of the block length is the same for both code rates. The curves are only shifted upwards as the code rate is increased. In order to reach the capacity limit a rather long block length (10^6) is required, but such block lengths are impractical for most systems. Typical block lengths can be the ATM cell size (424 bit) or MPEG packet size (1504 bit), and if we assume 1/2-rate coding this gives a capacity loss of 1.4 dB and 0.8 dB respectively.

2.5 How to Utilise the Limits

Code Design

Turbo codes have been made, which are very close to the Shannon Limit at moderate BER levels. These codes and other related codes are not algebraic codes, but depend on some parameters, which are found by a cut and try process. When

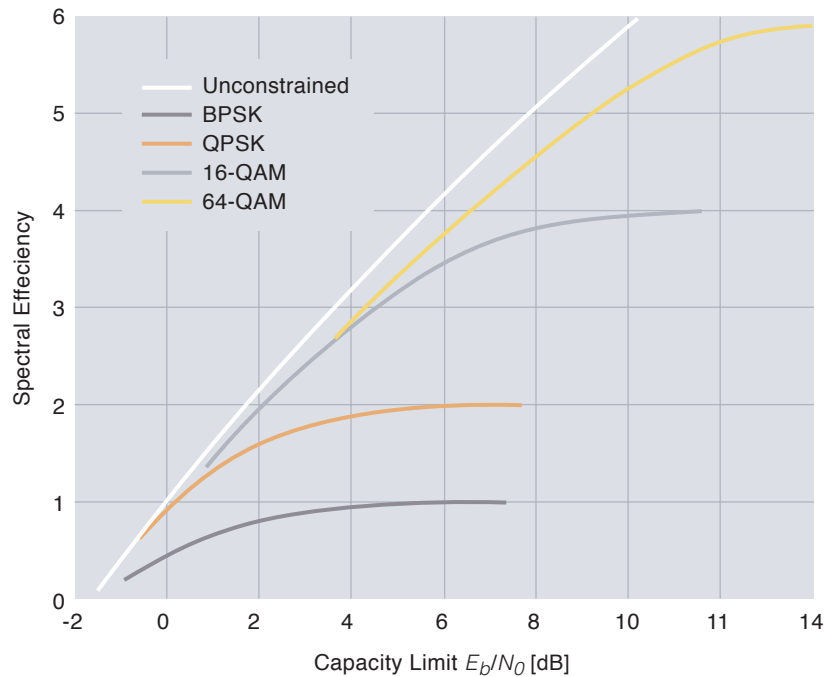


Figure 5 Spectral Efficiency as function capacity limit QAM

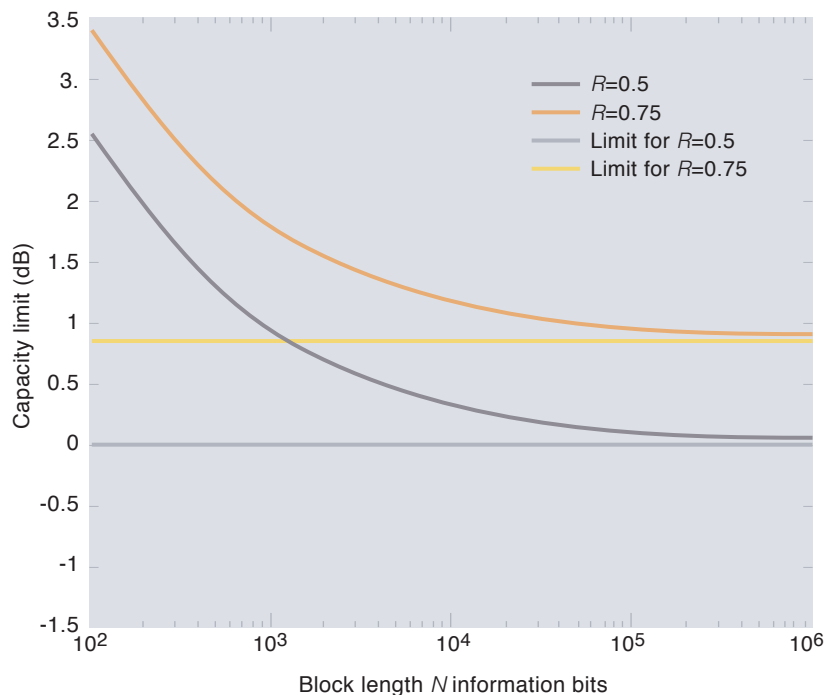


Figure 6 Capacity limit as function of block length (information bits)

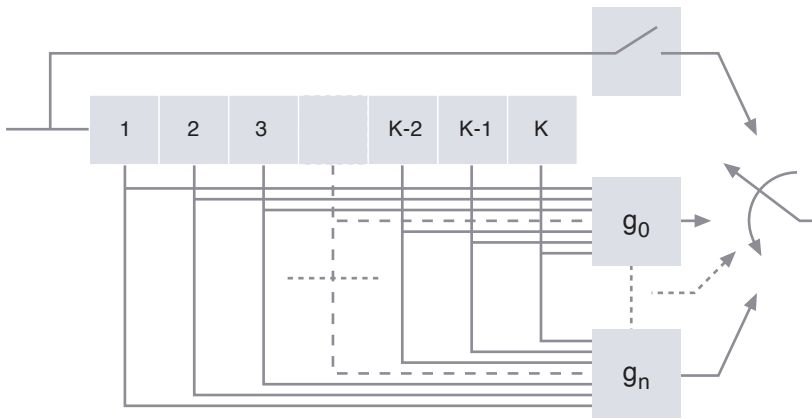


Figure 7 Convolutional feed forward encoder

doing simulation for a set of code parameters we can easily determine whether a code is good or not by comparing the simulation results against the theoretical limits. Further in this article we will use the limits to compare different coding schemes, since not all codes are directly comparable to each other (different code rate, block length).

System Design Level

New codes like Turbo codes offer very large flexibility with respect to code rates and block lengths. When evaluating different coding techniques to be used in a new system this usually includes extensive time-consuming simulations. Instead we can use the capacity limits to estimate the performance since the limits presented offer an easy way to get a first estimate.

3 Channel Coding for Radio Systems

3.1 Introduction

Since the radio channel is subject to various impairments like fading, interference and thermal noise, advanced channel coding is necessary for sufficient performance and range. In addition to the performance criteria, the coding scheme must also allow sufficient system flexibility. For satellite communication systems it might also be necessary to apply adaptive modulation and coding to account for varying quality of service requirements and channel variations. The coding scheme should then be able to provide variable error protection depending on the channel conditions. For fixed radio systems reliability and extremely low error rates are vital, and the flexibility is not easily utilised.

3.2 Coding for Satellite Applications

3.2.1 Viterbi Decoding of Convolutional Codes

Convolutional codes have been very popular for forward error correction in satellite systems. For convolutional codes, dependence between the symbols is obtained by performing a convolution on the data symbols. With more than one such linear combination, redundancy is added, and the code can correct errors. Figure 7 shows a feed forward convolutional encoder with constraint length K and code rate $R = 1/n$ or $R = 1/(n+1)$ if the systematic²⁾ encoder switch is closed. Which shift register contents to add (modulo 2) is decided by the generator polynomials, g_i . The outputs are then multiplexed onto the output line.

A nice feature with convolutional codes is the existence of a maximum likelihood sequence decoding algorithm originally proposed by Viterbi in 1967 [6] and now known as the Viterbi algorithm. The *constraint length*, K , of a convolutional code is commonly defined as the number of encoder memory (or delay) elements plus one³⁾. Since the complexity of the Viterbi algorithm increases exponentially with the memory of the encoder, we normally have to choose constraint lengths around 10 or lower to be able to decode with the Viterbi algorithm. The code rate, R , is the relation between the number of information bits, k , and code bits, n , transmitted, such that $R = k/n$. Rate k/n codes where k is different from one can be constructed by applying multiple shift registers. A problem with this approach is that the number of trellis states becomes 2^{kK} . Instead, high rate convolutional codes can be obtained by puncturing. Puncturing means that we delete some code bits of a lower rate $1/n$ code to obtain a code with a higher rate. For these codes the complexity of the decoding is essentially the same as for a rate $1/n$ code. Results have shown that the performance loss is also quite low by constructing families of multi-rate and rate compatible codes by puncturing [7] [8]. An important parameter for good performance of convolutional codes is the *free distance*, d_f . The free distance is the distance (or weight) of the path with the shortest distance from the correct path in the trellis.

A convolutional code that has become almost a standard for satellite communications is the rate $R = 1/2$ constraint length $K = 7$ convolutional code with generator polynomials 133 and 171 (in octal notation). It is for instance used in a number of Inmarsat systems, and in the Digital Video Broadcasting (DVB) satellite system. The

²⁾ When the encoder is systematic the information sequence appears directly in the coded sequence.

³⁾ Other definitions also exist in the literature.

Code rate	E_b / N_0 to achieve Performance at 10^{-6}	Capacity for constrained binary input	Capacity limit block length	Loss compared to constrained input limit
1/2	5.0 dB	0.2 dB	Not applicable	4.8 dB
3/4	6.0 dB	1.6 dB	Not applicable	4.4 dB

free distance of this code is 10. This is the highest possible free distance of a rate 1/2 feed-forward convolutional code with constraint length 7. This code turns out also to be optimum with regard to a more sophisticated criterion [9]. In Table 1 we give the required E_b / N_0 to achieve a bit error rate of 10^{-6} for this constraint length $K = 7$ code. Observe that although the code has a rather short constraint length, the performance is less than 5 dB away from the theoretical limit on a BPSK/QPSK channel.

3.2.2 Sequential Decoding of Convolutional Codes

As described above the complexity of the Viterbi decoding algorithm increases exponentially with the constraint length. At the same time the error probability of the convolutional code also decreases exponentially with the constraint length. To have sufficiently low bit error rates at low signal-to-noise ratios, we might have to use a constraint length that makes Viterbi decoding impractical. A possible solution is then to use sequential decoding, see for instance [10] or [11] for details. Sequential decoding is a sub-optimal decoding algorithm, but since the decoding complexity increases linearly with the constraint length instead of exponentially, we can always choose a constraint length that is sufficiently high to meet our *BER* requirements even if the decoding is sub-optimal.

The basic idea of sequential decoding is to search the code tree sequentially, going along the branch which gave the best metric increment. When the channels are good this is most likely the path that would have been chosen also by the Viterbi algorithm. When the channel is noisier, we will from time to time make an incorrect local decision in our search. This will usually be realised quite soon since the accumulated metric will now be bad. The decoder will thus back up and try alternative paths, which locally gave a worse metric. Obviously, when the channel is very noisy there will be a lot of back- and forward searches before the correct path is found.

Therefore, the decoding time will vary with the channel conditions and is thus a random variable. To cope with this varying decoding complexity, large buffers are required to store data in periods of intense search activity. With finite buffer sizes there is always a certain probability that the decoder has not finished decoding before the buffers are full (buffer overflow situation). The decoding must then be stopped. This is normally the most critical event with sequential decoding. Since the undetected error probability can be made sufficiently low by choosing a high constraint length, the overflow rate will dominate the total error rate. For sequential decoding the speed of the decoder will therefore have influence on the system performance. A better implementation or a faster hardware or signal processor will improve the error rate since the decoder has time for more search, and buffer overflows become more rare. To be able to recover quickly from an overflow situation, systematic encoders are often used with sequential decoding. It can be shown that if a channel/modulation parameter called the Pareto exponent is above one, $\rho > 1$, then the mean value of the number of computations required will be bounded. In this case a sufficiently large constraint length can be chosen to have negligible error rate.

“Inmarsat B High speed data” is a mobile/portable satellite service launched some years ago. This system uses a rate 1/2 systematic convolutional code with constraint length $K = 36$. Due to the very high constraint length, Viterbi decoding is impossible and sequential decoding is applied. The generator polynomial of the encoder is 714461625313 (octal notation). This code is an *optimum distance profile* (ODP) convolutional code [12], which makes it well suited for sequential decoding.

There has also been some work on sequential decoding for Rayleigh fading channels [13], and for such channels a higher signal-to-noise ratio is needed for the Pareto exponent to be one.

Table 1 Performance of constraint length 7 convolutional code with Viterbi decoding, compared to capacity limits

Table 2 Performance of sequential decoding, compared to capacity limits

Code rate	E_b / N_0 where Pareto exponent is equal to 1	Capacity for constrained binary input	Capacity limit block length	Loss compared to constrained input limit
1/2	2.2 dB	0.2 dB	Not applicable	2.0 dB
3/4	3.7 dB	1.6 dB	Not applicable	2.1 dB

Code	Code rate	E_b / N_0 to achieve BER of 10^{-5}	Capacity for constrained input	Capacity limit block length ⁴⁾	Loss at 10^{-5}
Voyager	0.44	2.5 dB	-0.1 dB	0.2 dB	2.3 dB
DVB-RCS	0.46	3.5 dB	0.0 dB	0.8 dB	2.7 dB

Table 3 Performance of Concatenated Codes (Convolutional and Reed-Solomon)

3.2.3 Concatenated Coding

Concatenated coding is a combination of an inner code and an outer code where the inner code should work well at low signal-to-noise ratios using maximum likelihood decoding (soft decoding) and a strong algebraic outer code with low redundancy. A much used combination is to use a convolutional code ($K = 7$) with Viterbi decoding and a Reed-Solomon outer code. The Reed-Solomon code is non-binary and typically works on 8-bit symbols. The errors from the Viterbi decoder are bursty and in order to combat these errors a byte interleaver is used to split up the long error burst. However, for packet transmission and time critical services interleaving cannot be used.

This coding scheme was for a long period the “state-of-the-art” in communication systems and has been used in deep-space communication (“Voyager” [14]) and is used in the DVB satellite system. The Voyager code uses a (255,223) Reed-Solomon code and DVB-RCS a (204,188) Reed-Solomon code [15]. In the DVB-RCS system one is limited to MPEG packets and cannot use interleaving, while for Voyager one could interleave up to 8 Reed-Solomon blocks. The performance is shown in Table 3.

3.2.4 Turbo Codes

Berrou et al. [16] introduced a new class of error correcting codes, called Turbo Codes. This coding technique is based on Parallel Concatenated Convolutional (PCC) codes, where two parallel

convolutional codes are separated by an interleaver, see Figure 8. The component encoders are RSC (Recursive Systematic Convolutional) codes. These encoders can be very short. Probably the most important element in the code is the interleaver, which defines the minimum distance. Designing the interleaver is critical for the performance, and the interleaver should ideally be random or pseudo-random [17]. The Turbo code can easily be applied to any block size and code rate, [16] and [18]. However one must be especially careful with the interleaver and puncturing map for high code rates [19].

The decoding is iterative and based on soft-input and soft-output.

To achieve this, the Turbo decoding algorithm may apply the Bahl-Jelinek-Cocke-Raviv (BJCR) algorithm [20], or the less optimum soft-output Viterbi algorithm (SOVA) [21]. A simplification of the BJCR algorithm has been shown to achieve very good results with very little performance loss [22], however several simplified versions exist. The actual complexity depends on component codes and the number of iterations required. The Turbo coding/decoding technique has achieved results very close to the Shannon Limit.

A lot of work has been done for PCC codes and has been proposed for a wide range of standards (EDGE, UMTS, DVB-RCS, CCSDS and several more). Commercial implementations can be

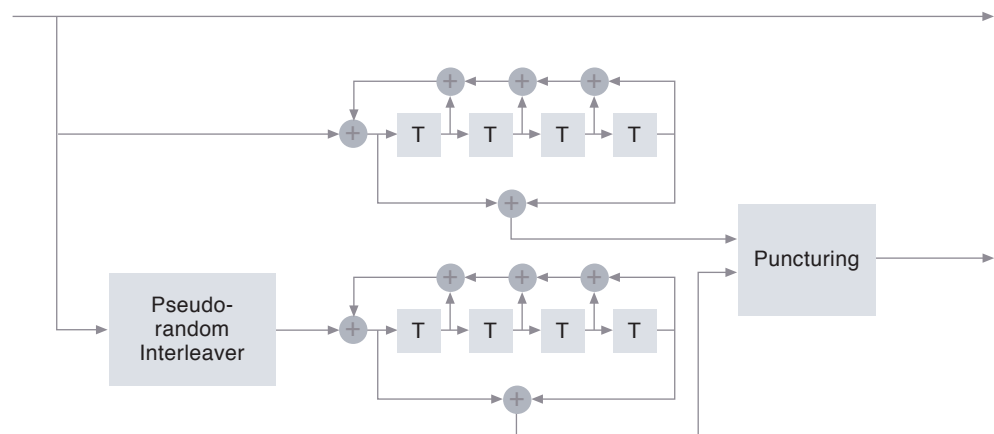


Figure 8 The original Turbo Code as presented by Berrou

⁴⁾ The block length for the Voyager code is based on 8 interleaved Reed-Solomon blocks while the DVB-RCS code does not use interleaving.

Code	Code rate	E_b / N_0 to achieve BER of 10^{-6}	Capacity for Constrained input	Capacity limit Block length ⁵⁾	Loss at 10^{-6}
Berrou code (variant)	1/2	1.6 dB	0.2 dB	1.0 dB	0.6
DVB-RCS Code	1/2	1.8 dB	0.2 dB	1.0 dB	0.8

bought, even though most of them are based on Field-Programmable Gate Array (FPGA) solutions, like the Nera Turbo decoder [23] and others [24], [25].

Nera uses both a variant of the original Berrou code and a similar but somewhat less complex Turbo code in its satellite products. The latter is used in the DVB Return Channel System (DVB-RCS), [15] [26]. In Table 4 the performance of these two codes is given for MPEG block sizes (188 bytes). We see that both codes are within 1 dB of the capacity limit.

The disadvantage of the Turbo code is that one might get a flattening of the bit error curve. However, with a good interleaver design this flattening effect can be mitigated and only occur at very low error rates. The inventors have patented both the Turbo encoding scheme and different decoders.

3.3 Comparison of Coding Techniques

In Figure 9 we compare convolutional coding and concatenated coding with Turbo together with some coding limits. The comparison is done for the MPEG packet size and approximately 1/2-rate coding. We see that all codes achieve a very good coding gain, but the Turbo code gives us an additional 2 dB compared to more traditional coding techniques. Appropriate simulation results for sequential decoding were not available, but results from [13] for a Rayleigh fading channel and a constraint length $K = 36$ code, show that $BER = 10^{-6}$ is achieved approximately 1 dB from the theoretical sequential limit. The difference is probably less for an AWGN channel. The sequential limit that is plotted is for 1/2-rate coding, corresponding to a Pareto exponent $\rho = 1$.

3.4 Coding for Systems with High Spectral Efficiency

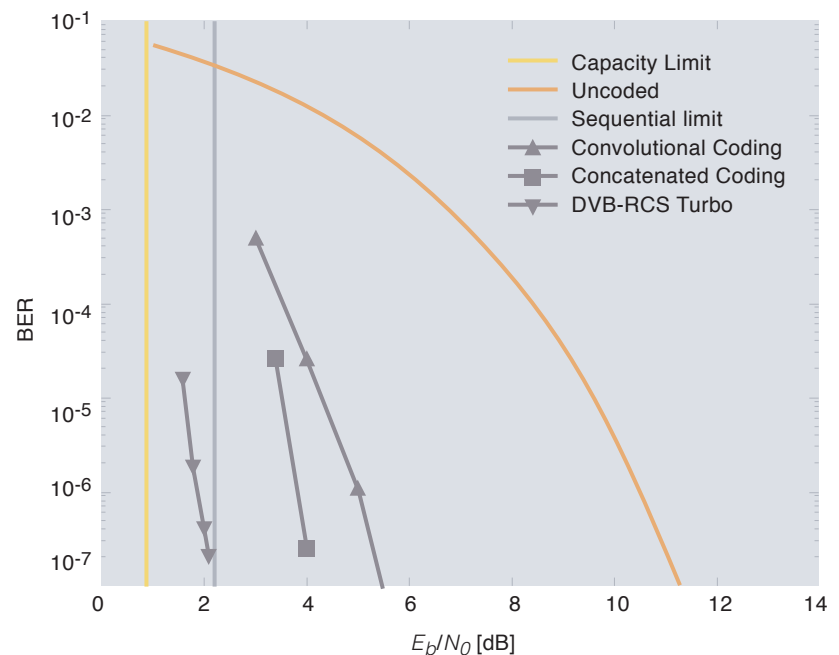
3.4.1 System Requirements

The most effective transmission medium both in terms of reliability and transmission speed is optical fibre transmission. However, in many

cases it is not possible or cost effective to apply fibre cables. The obvious, and often the only alternative is then radio. When the high data rates that are normally transmitted on fibre are to be transmitted via a radio system, we need modulation schemes with very high spectral efficiency, and coding schemes that can provide virtually error-free communication. The high spectral efficiency requirement means that the code rate must be rather high, otherwise the modulation alphabet must be extremely large. Virtually error-free communication means that the bit error rate must be below 10^{-12} . The last few years coding schemes have been found that are quite close to the Shannon limit⁶⁾. However, these schemes often also have rather large block lengths. Due to attenuation, fading and atmospheric disturbances the range of such a radio transmission scheme is limited to between 50 and 150 km where the specific range depends on the climatic conditions of the installation site, the topography and the carrier frequency that is applied. At high carrier frequencies the attenuation will be the limiting factor, while at lower frequencies multipath propagation may be the most critical. Therefore, in order to connect two

Table 4 Performance of the original Turbo and DVB-RCS code

Figure 9 Comparison of coding techniques



⁵⁾ The capacity of 1/2-rate coding with binary constrained input is 0.2, but including 0.8 loss due to the limited block size (MPEG) we get 1.0 dB.

⁶⁾ The distance to the Shannon limit is still higher for higher order modulation than for binary coding. Some improvement is expected by spectral shaping.

⁷⁾ We have used the term effective block length to indicate that it is not necessarily a block code that is applied, but for trellis codes the "block length" is not easily defined.

Table 5 Necessary spectral efficiency STM-1 in different channels

Channel Spacing	Practical Symbol Rate	Spectral Efficiency	Minimum modulation alphabet (QAM)
28 MHz	24	6.5	128-QAM
40 MHz	33	4.7	32-QAM
56 MHz	47	3.3	16-QAM

Spectral efficiency	Modulation alphabet	Minimum Code rate
3.3	16QAM	0.83
3.3	32QAM	0.66
3.3	64QAM	0.55
4.7	32QAM	0.94
4.7	64QAM	0.78
4.7	256QAM	0.58
6.5	64QAM	Not possible
6.5	128QAM	0.93
6.5	256QAM	0.81
6.5	1024QAM	0.65

Table 6 Necessary minimum code rate for a given spectral efficiency and modulation type

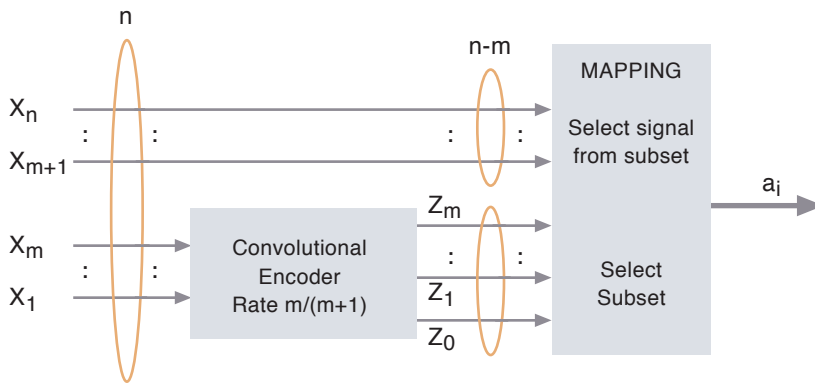


Figure 10 Trellis Coded Modulation (TCM) and mapping of bits to symbols

distant locations, we need several radio hops. Thus, to avoid a much higher delay than what is present for the fibre system, we also need to limit the processing delay for the radio transmitter and receiver. This puts strong requirements on the allowed effective block length⁷⁾ of the coding scheme.

Fixed radio systems are normally connected to the backbone network and must therefore at least offer STM-1 rates (155.52 Mbit/s). In most European countries these links must follow a frequency plan with a channel spacing of 28 MHz, 40 MHz or 56 MHz. In Table 5 we give the necessary spectral efficiencies in these cases. It would be beneficial to offer two STM-1 carriers in one channel or even to go to the higher STM-4 rate (622.02 Mbit/s), and to offer as high capacity and spectral efficiency as possible. However, implementation issues limit the possible spectral efficiency upward to around 8 bits/s/Hz.

Table 6 summarises the minimum code rates that can be applied in order to achieve a given spectral efficiency as presented in Table 5 for a given modulation method. We see that if a code rate close to 1/2 is to be used, we need a very large modulation alphabet to achieve sufficiently high spectral efficiency. Such large constellations will have other impacts on the system like problems with synchronisation and non-linearity effects in power amplifier and RF front end. Although we may obtain a spectral efficiency of 6.5 with a code rate 1/2 using 8192QAM, such a radio link will be extremely expensive or difficult to implement with existing technology.

Table 7 Performance and parameters for some specific Nera TCM schemes implemented in radio relay systems, XD means X dimensional

Parameter	Overall Code rate	Spectral Efficiency	Measurements E_b / N_0 at BER = 10^{-6}	Loss compared to unconstrained capacity limit
2D TCM, 32 QAM	0.80	4	11.7 dB	6.0 dB
2D TCM, 64QAM	0.83	5	14.0 dB	6.1 dB ⁸⁾
4D TCM, 128QAM	0.93	6.5	15.7 dB	4.3 dB

⁸⁾ Compared against the constrained input the loss is 4.8 dB.

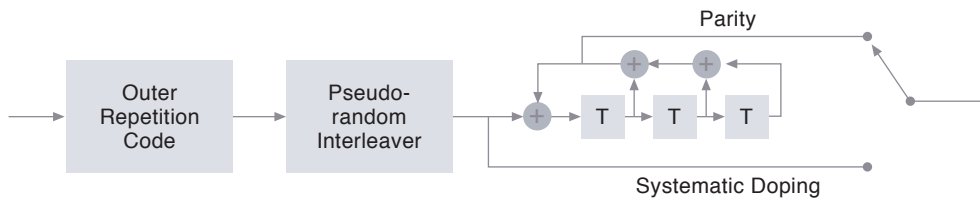


Figure 11 Serial Concatenated Convolutional code, ten Brink

3.4.2 Trellis Coded Modulation

Since Ungerboeck published the idea of trellis coded modulation (TCM) [28] combining the coding and modulation operations, TCM has become the obvious solution for spectrally efficient systems like radio relay systems. The codes normally used with TCM are rather short, and thus the delay does not lead to any serious problems. However, the mapping to the constellation points is a critical operation in order to achieve good performance. Figure 10 shows the basic idea of trellis coded modulation including the mapping operation to constellation symbols. Table 7 presents some performance results for TCM schemes used with radio links.

3.5 Promising New Coding Schemes

The coding schemes described so far have been applied for some time and are in practical use in various systems or prototypes. In this section we describe coding schemes that have so far mostly been subject to research, but are promising candidates with potential for future applications. The first is a serial concatenation of convolutional codes with iterative decoding. Then we look at a class of Turbo codes that are based on product codes (see for instance [29]), which are decoded iteratively. These codes have therefore been named Turbo Product Codes. The third method is a type of codes which was first proposed by Gallager as early as 1963 [30], and is known as Low Density Parity Check (LDPC) codes.

3.5.1 Serial Concatenated Convolutional Codes

Serial Concatenated Convolutional Codes (SCCC) was introduced by Benedetto et al. as an alternative to Turbo Codes based on serial concatenation [31]. Basically, the algorithms and much of the theory used for PCC can be applied for these codes. Note however that interleaver and component codes are different. According to [31] the outer code should be a maximum free distance, non-systematic convolutional code (NSC), while the inner code should be recursive. However, a somewhat different design is to use a simple repetition code ($R = 1/2$) as the outer code and rate 1 recursive convolution code as an inner code as proposed by ten Brink [32]. A

switch is used to replace some of the parity bits with systematic bits at a very low rate (1:100). This doping is necessary to “kick-start” the iterative decoder. The code has very low complexity, but still gets within 0.1 dB of the Shannon limit at 10^{-5} for very long block lengths.

Bit error rate (BER) performance results have shown these codes to be asymptotically better than PCC. There are few or no basic patents due to the many publications, especially by Benedetto. This scheme has slower convergence compared to PCC, which might stem from the fact that the inner and outer decoder will work at different signal to noise ratios. Good high code rate performance seems harder to achieve. It is unclear whether this is a fundamental problem or just lack of good codes.

3.5.2 Turbo Product Codes

The idea of Turbo Product Codes (TPC) is to apply iterative (or Turbo) decoding to the decoding of two concatenated block codes, known as product codes (see for instance [29]). In Figure 12 we show the fundamental idea of Product Codes. A number of code words are generated using the first component code (the row vectors), and then encoding is performed “vertically” by the second component code. Observe that we also encode the parity symbols creating parity on the parity. The two codes that are applied may be different or the same. Typically BCH codes, Hamming codes and/or parity check codes are used. We may also apply a third code, requiring a cube for illustration. To be able to perform iterative decoding, we need soft decisions both in and out of the individual decoding processes. Details of this decoder can be found in the paper [33] by Phylindia et al. who first published the idea of Turbo Product Codes. These codes have since gained much interest due to some attractive properties. Block codes typically perform better than convolutional codes for high code rates. Turbo Product Codes therefore seem to perform better than the convolutional code based Turbo codes for high code rates. Therefore, we consider Turbo product codes to be a strong candidate for replacing TCM in the high spectral efficiency radio systems described above. Also,

⁹⁾ This feature is sometimes erroneously referred to as an error floor. We prefer to call it a levelling out feature since it is certainly not a floor, and the decrease in error rate is still better than for the uncoded system.

Figure 12 The principle of product codes: parity symbols are calculated for a data vector, and then vertically across the previously calculated code words with the second code

d ₁₁	d ₁₂	d ₁₃	d ₁₄	d ₁₅	P ₁₆	P ₁₇
d ₂₁	d ₂₂	d ₂₃	d ₂₄	d ₂₅	P ₂₆	P ₂₇
d ₃₁	d ₃₂	d ₃₃	d ₃₄	d ₃₅	P ₃₆	P ₃₇
P ₄₁	P ₄₂	P ₄₃	P ₄₄	P ₄₅	P ₄₆	P ₄₇

the levelling out effect⁹⁾ of the bit error rate curve that has been observed for other Turbo codes seem less dramatic for the product codes than the convolutional codes. There are indications that the levelling out effect can be made negligible for many applications by some smart modifications [34]. For the PCC codes this effect is to some extent controlled or reduced by careful interleaver design, improving the minimum distance of the code.

As described, product codes consist of two or more concatenated block codes as component codes. Therefore, the Turbo Product Codes have the same restriction on the relation between code rate and block length as their component codes. Thus, we cannot get any arbitrary code rate matching any desired block length. The total block length of the product codes is the product of the block length of each component code. This means also that the code rate is given by

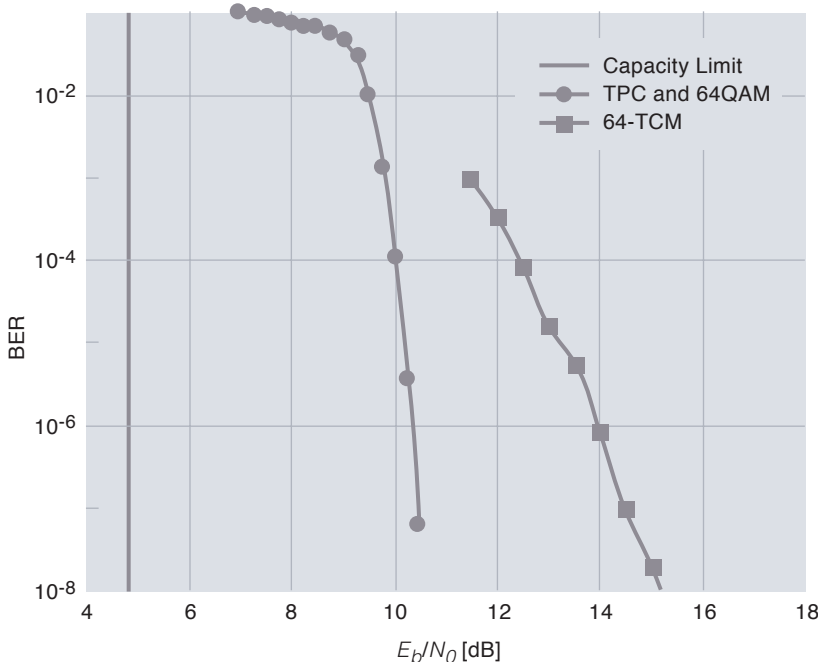
$$R = \frac{k_1 k_2}{n_1 n_2},$$

where k_i is the number of information bits encoded into a block with code i , and n_i is the total number of encoded bits in the code i . This lack of flexibility for TPC may be a rather undesired feature for many applications.

The authors are not aware of any commercial radio systems in operation today using Turbo Product Codes. However, these codes were seriously considered for several satellite systems that were standardised or planned recently. TPC is also an option in the *IEEE* 802.16 standard for wireless metropolitan area networks that was ratified in December 2001. Since these codes now start to appear in standards, there are also chip manufacturers that have TPC ASICs under development.

Figure 13 shows the performance of TPC with 64QAM modulation plotted together with 64TCM. The spectral efficiency of the 64TCM scheme is 5, while the spectral efficiency of the 64QAM modulated TPC is 4.7 (the code rate used is $R = 0.779$, and two BCH (64,57) component codes are used). Observe that the spectral efficiency is slightly lower for the TPC scheme, but it also significantly outperforms the TCM scheme. Because of the length of the TPC code the delay is also somewhat higher than for TCM.

Figure 13 Turbo Product Codes with 64QAM compared to 64TCM



3.5.3 Low Density Parity Check Codes

Low Density Parity Check (LDPC) codes are block codes that were first presented by Gallager [30], and are also referred to as Gallager codes. These codes were “forgotten” for many years until they were rediscovered by McKay [35]. LDPC codes probably have the world record on coding gain with a performance only 0.005 dB from the Shannon limit [36] (using extremely long code words). These codes have a very sparse parity check matrix, H . A sparse parity matrix means that the number of ones is very low compared to the number of zeros. This is why the codes are referred to as low density parity check (LDPC) codes. The parity check matrix is constructed in a semi-random manner, placing t ones on each column. Research results [35] indicate that it is advantageous to have an overlap of only one (overlap constraint) between the ones of the columns¹⁰⁾. In addition we try to achieve as uniform a row weight as possible. A parity check matrix constructed as described above is then transformed into systematic form, and the generator matrix G is easily obtained and used for encoding. For a given G the encoding process is exactly as for any block code.

¹⁰⁾ This is to avoid short cycles in the graph, which are bad for the decoder performance.

Gallager codes, where the columns of the parity check matrix have the same (uniform) weight, are called regular Gallager codes. Similarly, irregular Gallager codes have parity check matrices with non-uniform column weights. There are indications that irregular Gallager codes have better performance than regular Gallager codes [35]. Gallager codes may also be constructed over $GF(q)$. Such non-binary codes may be better than the irregular Gallager codes over $GF(2)$ [35].

Gallager codes are decoded iteratively by a sum-product algorithm. After each iteration the syndrome is calculated and a new iteration is performed if the syndrome is not zero. When the syndrome is zero, we have found a valid code word. The decoding is then terminated and the information bits are released. If the syndrome is not zero after some predefined maximum number of iterations, the decoding is halted and the information part of the current code word estimate is released. The minimum distance of Gallager codes is high, and, except for very short block lengths, errors are normally only observed when the decoding is terminated because the maximum number of iterations was reached. Thus, Gallager codes have a built-in error detection mechanism that can make them promising candidates for packet data communication with retransmission (hybrid ARQ). Gallager codes need many more iterations in the decoding than the Turbo codes, but the complexity of each iteration is much lower.

In contrast to the decoders for many other block codes, the decoder for Gallager codes can easily utilise soft decisions. In the decoder the soft decisions are used to initialise to appropriate initial symbol probabilities. Consequently, there is also limited extra complexity in soft decision decoding of Gallager codes. Figure 14 shows the performance of some LDPC codes with MPEG and ATM packet sizes. The performance is comparable to or better than the performance of the best Turbo codes.

4 Comparison and Discussion

As a comparison of the discussed coding schemes, we have plotted the spectral efficiency of a number of coding techniques as a function of the signal-to-noise ratio (E_b / N_0). In Figure 15 we present the spectral efficiency for ATM sized (i.e. 53 bytes or 424 information bits) packets as a function of the required E_b / N_0 to achieve a packet error rate equal to 10^{-5} , where E_b is the bit energy and N_0 the noise spectral density. Note that for the LDPC codes we have used bit error rate instead of packet error rate, and therefore the results are slightly too good compared to the other schemes. However, due to the steep curves the SNR difference resulting from this is modest. We see that the iterative (or

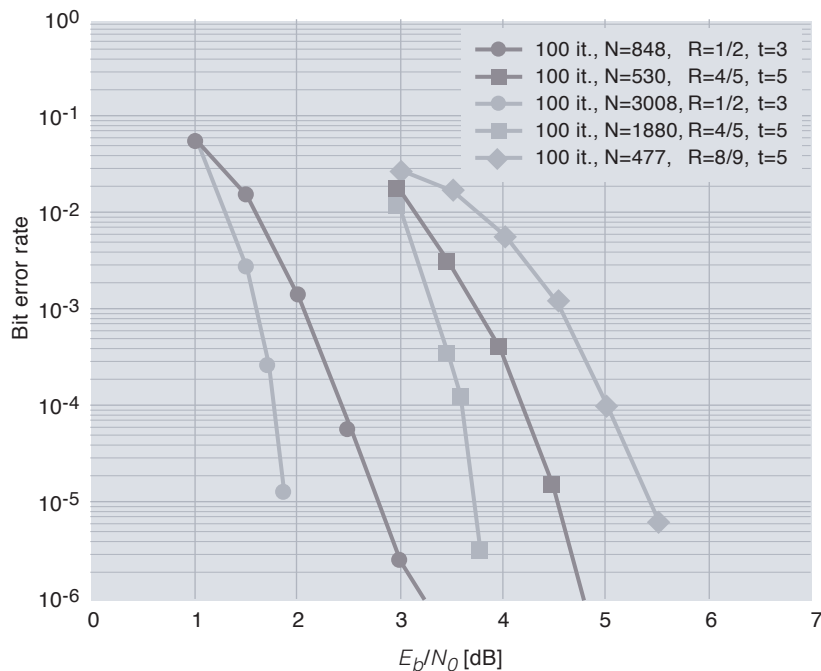


Figure 14 Performance of LDPC codes with various packet sizes (MPEG and ATM packet sizes) and various code rates. Results are shown for regular LDPC codes with t denoting the column weight of the parity check matrix

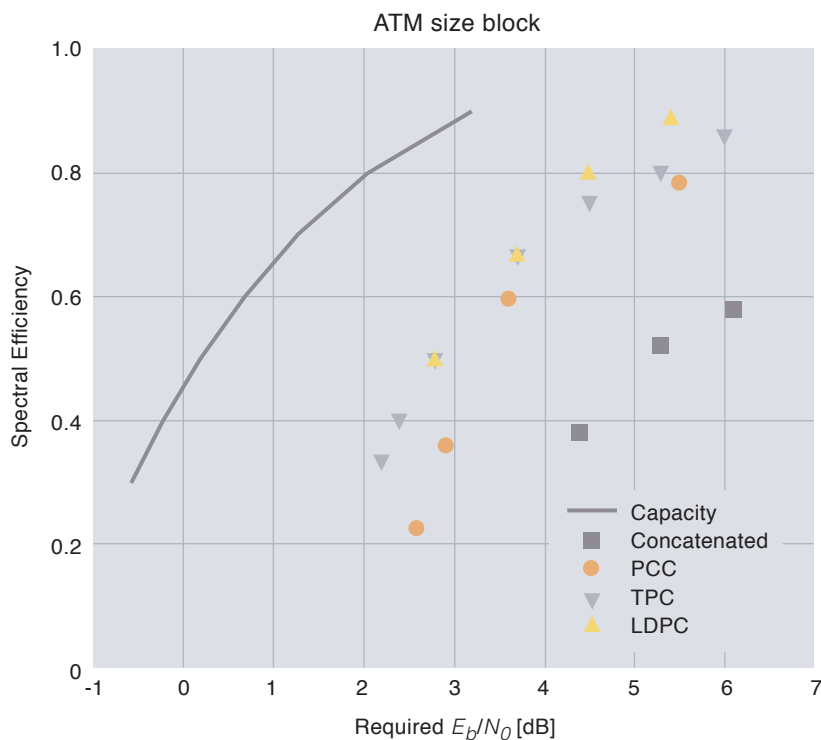


Figure 15 Spectral efficiency of various error control schemes for ATM sized packets. We have applied a packet error rate of 10^{-5} , except for LDPC codes a bit error rate equal to 10^{-5} is applied. Modulation is BPSK

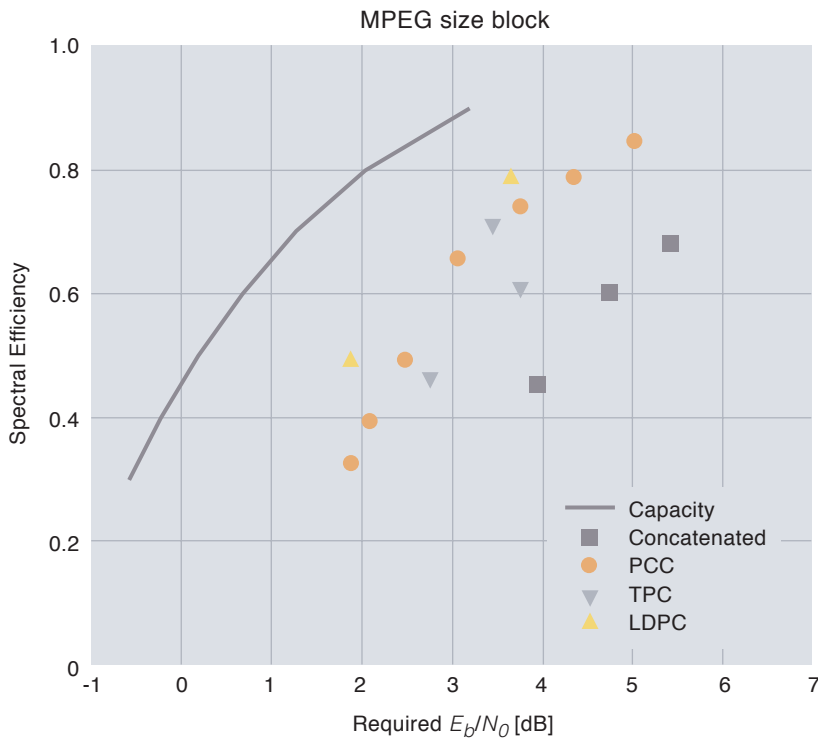


Figure 16 Spectral efficiency of various error control schemes for MPEG sized packets. We have applied a packet error rate of 10^{-5} , except for LDPC codes a bit error rate equal to 10^{-5} is applied. Modulation is BPSK

Turbo based) schemes perform significantly better than the concatenated schemes. In Figure 16 we have plotted the corresponding results for MPEG packets (i.e. 188 bytes or 1504 information bits). The concatenated coding scheme does not apply soft decisions when decoding the Reed Solomon code since this would be rather complex. Soft decisions explain a major part of the performance difference.

The coding schemes presented also have different degrees of flexibility. The PCC-Turbo code can accommodate any block length by changing the interleaver size. We can change the code rate by altering the puncturing pattern. Thus, code rate and block length can be modified independently of each other. This gives great flexibility in the system design. For the Turbo Product Codes the code rates and block lengths are directly given by the code rate and block length of the component codes. Shortening of the code is of course possible, but still we do not have the same flexibility as for the PCC Turbo codes. Note the relatively poor performance of the $R = 0.6$ code for TPC and MPEG packets. This comes from the limited flexibility of the TPC code that is applied. With additional component codes one could achieve improved performance. The LDPC codes are somewhere between these two other Turbo codes in flexibility. A code with any code rate and block length can in principle be designed, but distinct encoders and decoders must be implemented. There are indications of significant loss when puncturing the LDPC codes [37], so changing the code rate by puncturing is not recommended.

References

- 1 Sheriff, R E, Hu, Y F. *Mobile satellite communication networks*. West Sussex, John Wiley, 2001.
- 2 Shannon, C E. Communication in the Presence of Noise. In: *Proceeding IRE*, 37, 10–21, January 1949. (Reprint available in *Proceedings of the IEEE*, 86 (2), 447–457, 1998.)
- 3 Blahut, R E. *Principles and Practise of Information Theory*. Addison-Wesley, 1987.
- 4 Proakis, J G. *Digital Communications*. 2nd Edition. New York, McGrawHill, 1989.
- 5 Dolinar, S, Divsalar, D, Pollara, F. *Code Performance as a Function of Block Size*. Pasadena, California, Jet Propulsion Laboratory, California Institute of Technology, 1998. (TMO Progress Report 42-133.)
- 6 Viterbi, A J. Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. *IEEE Transactions on Information Theory*, IT-13 (2), 260–269, 1967.
- 7 Frenger, P et al. Rate compatible convolutional codes for multirate DS-CDMA systems. In: *IEEE Transactions on Communications*, 47 (6), 828–836, 1999.
- 8 Frenger, P et al. *Multirate convolutional codes*. Gothenburg, Sweden, Chalmers University of Technology, Dept. of Signals and Systems, Communication Systems Group, 1998. (Tech. Rep. 21.)
- 9 Frenger, P, Orten, P, Ottosson, T. Convolutional codes with optimum distance spectrum. *IEEE Communication Letters*, 3 (11), 317–319, 1999.
- 10 Viterbi, A J. *Principles of digital communication and coding*. New York, McGraw-Hill, 1979.
- 11 Wicker, S B. *Error control systems for digital communication and storage*. New Jersey, Prentice Hall, 1995.
- 12 Johannesson, R. Some long rate one-half binary convolutional codes with an optimum distance profile. *IEEE Transactions on information Theory*, IT-22 (5), 629–631, 1976.
- 13 Orten, P, Svensson, A. Sequential decoding of convolutional codes for Rayleigh fading channels. *Wireless Personal Communica-*

- tions, 20 (1), 61–74, 2002. Kluwer Academic Publishers.
- 14 Costello, D J et al. Applications of Error-Control Coding. *IEEE Transactions on Information Theory*, 44 (6), 2531–2560, 1998.
 - 15 ETSI. *Standard Digital Video Broadcasting (DVB); Interaction channel for Satellite Distribution Systems*. Sophia Antipolis, 2000. (EN 301 790 V.1.2.2, 2000-12.)
 - 16 Glavieux, A, Berrou, C, Thitimajshima, P. Near Shannon Limit error correcting coding and decoding: Turbo-Codes (1). In: *Proceedings of IEEE International Conference on Communications*, Geneva, Switzerland, May 1993, 1064–1070.
 - 17 Crozier, S et al (eds.). Performance of Turbo Codes with Relative Prime and Golden Interleaving Strategies. *Sixth International Conference on Mobile Satellite Communication (IMSC '99)*, Ottawa, Canada, June 1999. (www.cra.ca/fec)
 - 18 Pyndiah, R et al. Near optimum decoding of products codes. In: *Proceedings of IEEE Global Telecommunications Conference (GLOBECOM '94)*, San Francisco, Nov–Dec 1994, 1/3, 339–343.
 - 19 Acikel, O F, Ryan, W E. Punctured turbo-codes for BPSK/QPSK channels. *IEEE Transactions on Communications*, COM-47 (9), 1315–1323, 1999.
 - 20 Bahl, L R et al. Optimal decoding of linear codes for minimizing symbol error rate. *Transactions on Information Theory*, IT-20, 284–287, 1974.
 - 21 Hagenauer, J, Offer, E, Papke, L. Iterative decoding of binary block and convolutional codes. *IEEE Trans. on Inform. Theory*, 42 (2), 429–445, 1996.
 - 22 Pietrobon, S, Barbulescu, A S. A simplification of the modified Bahl decoding algorithm for systematic convolutional codes. *Int. Symposium on Information Theory & its Applications*, Sydney, Australia, Nov. 1994, 1073–1077.
 - 23 Risløw, B et al. Implementation of a 64 kbit/s Satellite Modem Using Turbo Codes and 16-QAM. In: *Proceedings of Norwegian Signal Processing Symposium*, Tromsø, May 23–24, 1997.
 - 24 SmallWorld. (May 7, 2002) [online] – URL: <http://www.sworld.com.au/>
 - 25 TurboConcept. (May 7, 2002) [online] – URL: <http://www.turboconcept.com/>
 - 26 ETSI. *Digital Video Broadcasting (DVB); Interaction channel for Satellite Distribution Systems, Guideline of use of EN 301 790*. Sophia Antipolis, 2001. (ETSI Technical Report TR 101 790, 2001-06.)
 - 27 Forney, G D Jr, Ungerboeck, G. Modulation and Coding for Linear Gaussian Channels. *IEEE Transactions on Information Theory*, 44 (6), 2384–2415, 1998.
 - 28 Ungerboeck, G. Trellis-Coded Modulation with redundant signal sets. Part 1: Introduction. *IEEE Communications Magazine*, 25 (2), 5–11, 1987.
 - 29 Wilson, S. *Digital Modulation and coding*. New Jersey, Prentice Hall, 1996.
 - 30 Gallager, R G. *Low density parity-check codes*. Cambridge, MA, MIT Press, 1963. (Research Monograph series no 21.)
 - 31 Benedetto, S et al. Serial Concatenation of Interleaved Codes : Performance Analysis, Design and Iterative Decoding. *IEEE Transactions on Information Theory*, 44 (3), 909–926, 1998.
 - 32 ten Brink, S. Rate one-half code for approaching the Shannon limit by 0.1 dB. *IEE Electronics Letters*, 36 (1), 1293–1294, 2000.
 - 33 Pyndiah, R. Near optimum decoding of product codes: Block Turbo Codes. *IEEE Transactions on Communications*, 46 (8), 1003–1010, 1998.
 - 34 Advanced Hardware Architectures. (May 7, 2002) [online] – URL: <http://www.aha.com/>
 - 35 MacKay, D J. Good error-correcting codes based on very sparse matrices. *IEEE Transaction on Information Theory*, 45 (2), 399–431, 1999.
 - 36 Chung, S et al. On the design of low-density parity-check codes within 0.0045 dB of the Shannon limit. *IEEE Communications Letters*, 5 (2), 58–60, 2001.
 - 37 Orten, P. *Low density parity check codes for wireless communications*. Billingstad, Nera Research, 2000. (Technical Report.)

A New Look at the Exact BER Evaluation of PAM, QAM and PSK Constellations^{*})

PAVAN K. VITTHALADEVUNI AND MOHAMED-SLIM ALOUINI



Pavan K. Vitthaladevuni (23) received his B.Tech. degree in Electrical Engineering from the Indian Institute of Technology (IIT), Madras, India in 1999 and his MSEE degree from the University of Minnesota, Twin Cities in 2001. He is currently a graduate research assistant in the Department of Electrical and Computer Engineering at the University of Minnesota, Twin Cities and is working towards his PhD. His research interests span digital communications over wireless channels and information theory.

pavan@ece.umn.edu



Mohamed-Slim Alouini (33) received his Diplome d'Ingenieur degree from TELECOM Paris and his Diplome d'Etudes Approfondies in Electronics from the Univ. of Pierre & Marie Curie, Paris, both in 1993. He received his M.S.E.E. from Georgia Tech, Atlanta, in 1995, and his PhD in electrical engineering from Caltech, Pasadena, in 1998. He joined the department of Electrical and Computer Engineering of the Univ. of Minnesota in 1998, where his current research interests include statistical modeling of multipath fading channels, adaptive and hierarchical modulation techniques, diversity systems, interference mitigation techniques, and digital communication over fading channels. alouini@ece.umn.edu

Hierarchical constellations offer a different degree of protection to the transmitted messages according to their relative importance. As such they found interesting application in digital video broadcasting systems as well as wireless multimedia services. Although a great deal of attention has been devoted in the recent literature to come up with explicit closed-form expressions for the bit error rate (BER) performance of standard pulse amplitude modulation (PAM), quadrature amplitude modulation (QAM), and phase-shift-keying (PSK) constellations, very few results are known on the BER performance of hierarchical constellations. In this paper, we argue that a recursive way of Gray coding a constellation ensures the existence of a recursive algorithm for the exact and generic (in the constellation size) BER computation of generalized hierarchical PAM, QAM, and PSK constellations over additive white Gaussian channel (AWGN) channels. This new approach provides also as a byproduct an alternative unified way for the BER evaluation of the well-known standard PAM, QAM, and PSK constellations. Because of its generic nature, this new approach readily allows numerical evaluation for various cases of practical interest.

I. Introduction

In his study of broadcast channels, Cover [1] showed about three decades ago that one strategy to guarantee basic communication in all conditions is to divide the broadcasted messages into two or more classes and to give every class a different degree of protection according to its importance. The goal is that the most important information (known as basic or coarse data) must be recovered by all receivers while the less important information (known as refinement, detail, or enhancement data) can only be recovered by the "fortunate" receivers which benefit either from better propagation conditions (e.g. closer to the transmitter and/or with a direct line-of-sight path) or from better RF devices (e.g. lower noise amplifiers or higher antenna gains). Motivated by this information-theoretic study, many researchers have shown since then that one practical way of achieving this goal relies on the idea of hierarchical constellations (known also as embedded, multi-resolution, or asymmetric constellations) which consist of non-uniformly spaced signal points [2], [3], [4]. This concept was studied further in the early nineties for digital video broadcasting systems [3], [5] and has gained more recently new actuality with

- the demand to support multimedia services by simultaneous transmission of different types of traffic, each with its own quality requirement [6], [7], [8]; and
- a possible application in the DVB-T standard [9] in which hierarchical modulations can be used on OFDM subcarriers.

The evaluation of the bit error rate (BER) of classical uniform M -ary pulse amplitude modulation (PAM), M -ary quadrature amplitude modulation (QAM), and M -ary phase-shift-keyed (PSK) constellations has long been of interest. For example, exact expressions for the BER of 16-QAM and 64-QAM were derived in [10]. Later on, generic (in M) but approximate BER expressions for uniform M -QAM have been developed in [11] and [12]. More recently, Yoon et al. [13] obtained the explicit and generic (in M) closed-form expression for the BER of uniform square QAM. We extended these results to hierarchical square and non-square $4/M$ -QAM constellations (see [14]). For this particular family of hierarchical constellations, for every channel access, two bits are assigned for the basic information and $(\log_2 M - 2)$ bits are assigned for the refinement information. However, for more general hierarchical PAM, QAM or PSK constellations, the derivation of exact and generic closed-form expressions becomes more tedious. In this paper, we take a new look at this problem and we present an alternative recursive approach for the exact BER computation of generalized hierarchical M -PAM, M -QAM, and M -PSK constellations over additive white Gaussian channel (AWGN) channels. For example, we consider the 2/4-PAM wherein a BPSK constellation is embedded into a 4-PAM constellation (see Figure 1) and the 2/4/8-PAM constellation wherein a BPSK constellation is embedded into a 4-PAM constellation, which is in turn embedded into a 8-PAM constellation (see Figure 2). For these hierarchical constellations one information bit is sent for each of the two or three different levels

^{*}) This work is supported in part by the National Science Foundation (NSF) grant CCR-9983462 and in part by the Center of Transportation Studies (CTS) through the Intelligent Transport Systems (ITS) Institute, Minneapolis, Minnesota.

of protection required, respectively. Apart from solving for the exact BER of these generalized hierarchical constellations, this new approach provides as a byproduct an alternative unified way for the BER evaluation of the well-known standard uniform constellations.

The remainder of this paper is organized as follows. The subsequent section presents some design issues as well as the model and parameters of generalized hierarchical constellations. While section III describes the recursive algorithm for the exact BER computation of the constellations under consideration, section IV shows how these general results can be exploited to obtain the BER of standard uniform constellations. Section V illustrates the mathematical formalism by some numerical examples. Next, section VI outlines some potential applications of these new results. Finally, section VII concludes the paper with a summary of the main points of the paper.

II. System Model and Parameters

II.A General Setup

II.A.1 PAM

Consider a situation wherein we want to achieve unequal error protection for various bits in a PAM symbol. Given that information rides on the amplitude of the symbol, we need to construct a specific PAM constellation, to achieve the stated goal of unequal error protection. We consider a generalized $2/4/\dots/M$ -PAM constellation with Karnaugh map style Gray mapping (see Figure 2 for the generalized 8-PAM example). In the case of 8-PAM, we assume that there are 3 streams of data, each of which has a priority (or equivalently, a target BER). We take one bit from each stream to form a symbol of 3 bits. The bit from the bit stream with the highest priority (lowest allowed BER) is assigned to the

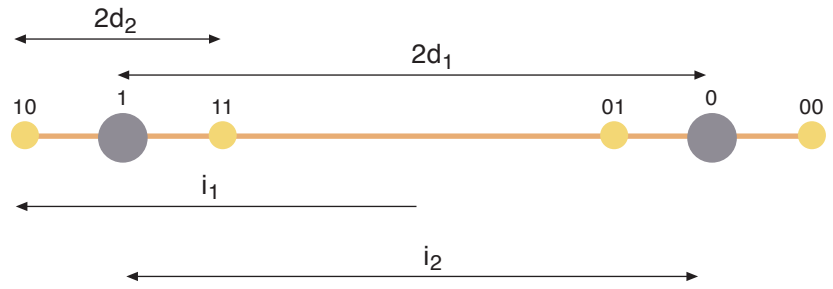


Figure 1 Generalized hierarchical 2/4-PAM constellation

most significant position, and as such, is referred to as the most significant bit (MSB). The bit from the bit stream with the second highest priority is assigned the next most significant position, and the bit with the least priority is assigned the least significant position, and is referred to as the least significant bit (LSB). In general, if we were to have m bit streams of data (referred to as sub-channels, i_1 through i_m , hereafter) with their respective priorities, we follow a similar procedure in assigning the bits to positions in the m -bit symbol. We would then have $M = 2^m$ symbols, and these can be visualized as points in an M -PAM constellation. In addition to arranging the bits in a symbol with respect to (w.r.t.) their priorities, we need to place these symbols in a hierarchical way in the constellation to achieve unequal error protection. In Figure 2, the fictitious BPSK constellation denoted by the large grey circle represents the highest priority (hereafter referred to as first level of hierarchy or sub-channel i_1). The distance d_1 refers to the highest priority. The distance d_2 represents the second priority (second level of hierarchy or sub-channel i_2). Finally, the distance d_3 represents the third priority (third level of hierarchy or sub-channel i_3). In other words, the constellation (A through H) can be visualized as a BPSK (S_1 and S_2) embedded into a 4-PAM (T_1 through T_4), which is further embedded into an 8-PAM (A through H). Gray coding is also done hierarchically. The fictitious BPSK is coded using 1

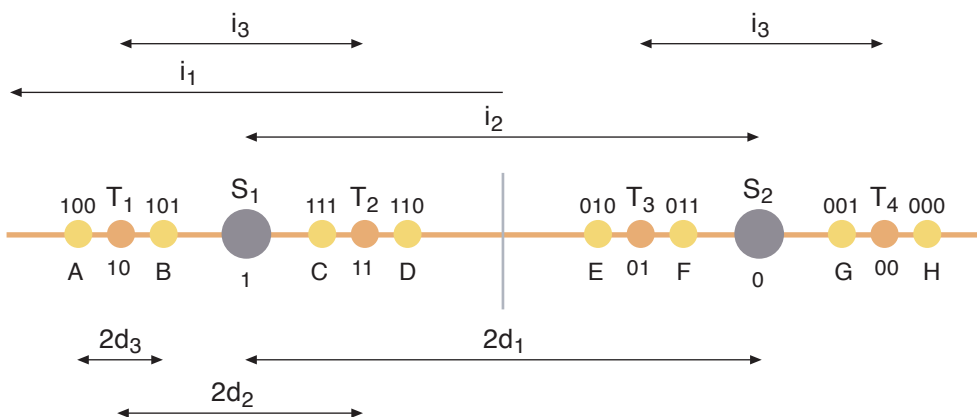


Figure 2 Generalized hierarchical 2/4/8-PAM constellations

bit. As shown in Figure 2, S_1 is coded as 1. In the next level of hierarchy, the fictitious T_1 and T_2 are coded as 10 and 11 respectively. Note that we are tagging a 0 or 1 to the right of the Gray code of S_1 . A and B are then coded as 100 and 101 respectively, by tagging a 0 and 1 to the right of the Gray code for T_1 . We remind the reader that A through H are the symbols that are transmitted. It helps to visualize the 8-PAM constellation as one that evolves from a 4-PAM which in turn evolves from a BPSK. This is why we will refer to this 8-PAM constellation as a 2/4/8-PAM. This procedure of Gray coding ensures the greatest protection to the MSB at the cost of the LSB.

In general, given any M -PAM ($M = 2^m$), this process can be automated as follows. Label the constellation points from the left to right (or vice versa) with integers starting from 0 to $M - 1$. Then, convert the integer labels to their binary form (for example, if we are dealing with a 16-PAM ($16 = 2^4$), the binary representation of 4 would be 0100). For the k -th symbol, let the binary equivalent be $b_{1,k} b_{2,k} \dots b_{m,k}$. Then, the corresponding Gray code ($g_{1,k} g_{2,k} \dots g_{m,k}$) ($k = 0, 1, 2, \dots, M - 1$) with integer label b ($b_{1,k} b_{2,k} \dots b_{m,k}$) is given by

$$\begin{aligned} g_{1,k} &= b_{1,k}, \\ g_{i,k} &= b_{i,k} \oplus b_{i-1,k}, \quad i = 2, 3, \dots, m, \end{aligned} \quad (1)$$

where \oplus represents modulo-2 addition. For reader convenience, MATLAB programs generating Gray codes are available at [15].

II.A.2 QAM

A generalized hierarchical square M -QAM constellation ($M = 2^{2m}$) can be modeled as follows. We assume that there are m bit streams of data

$$\left(m = \frac{1}{2} \log_2 M \right).$$

Each one of these incoming streams carries information of a particular priority. For every channel access 2 bits are chosen from each level of priority. The 2 highest priority bits are assigned the MSB positions in the in-phase(I) and the quadrature phase(Q), respectively. Bits with lower priorities are assigned the subsequent positions of lower significance. For instance the 2 bits with the second highest priority are assigned the second most significant positions in the I-phase and Q-phase, and so on until the 2 least priority bits are assigned the LSB position in the I- and Q-phase. This can be viewed as a 4/16/64.../ M -QAM constellation.

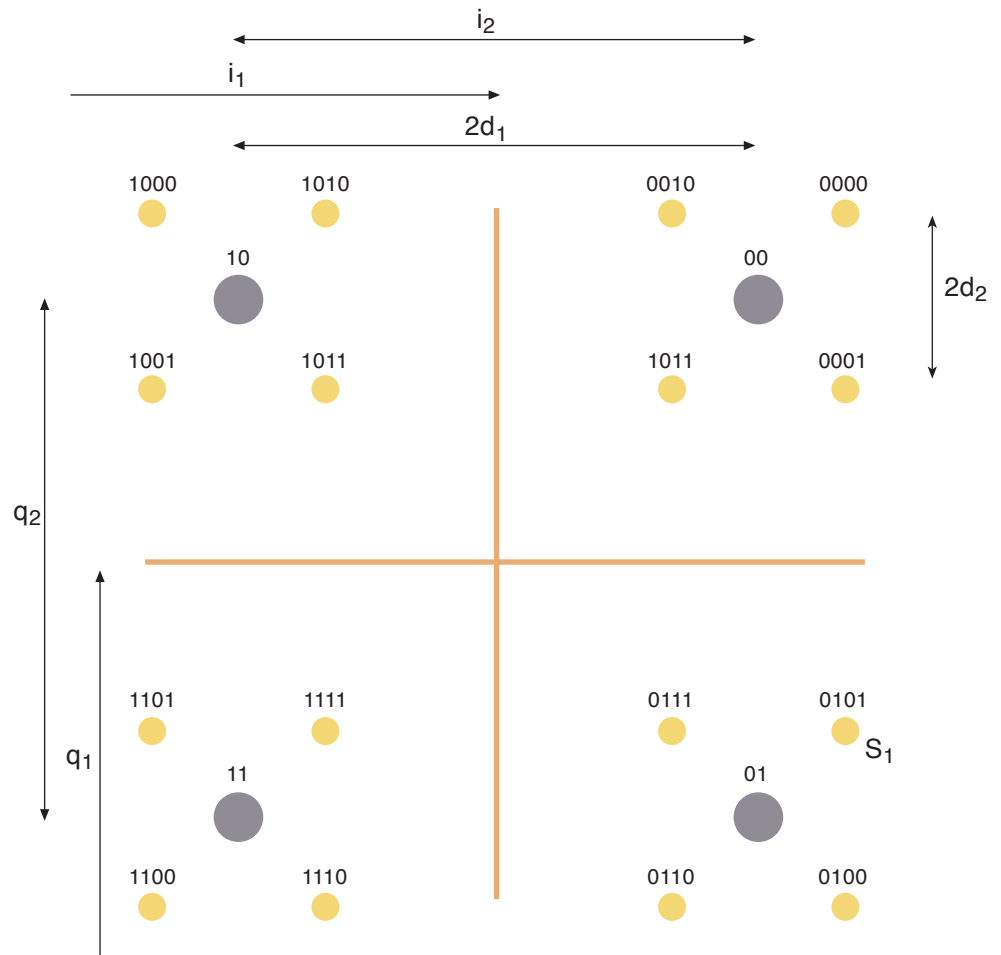


Figure 3 Generalized hierarchical 4/16-QAM constellations

The Gray codes for the symbols are given by $g_1^i g_1^q g_2^i g_2^q \cdots g_m^i g_m^q$, where the superscripts “ i ” and “ q ” refer to the in-phase and quadrature phase respectively. In other words, the Gray code for the symbol is obtained by interleaving the Gray codes of the symbol position in the I-channel and Q-channel \sqrt{M} PAMs respectively. For instance, in the 4/16-QAM constellation shown in Figure 3, the symbol S_1 has the I-channel Gray code as 00 (rightmost), and the Q-channel Gray code as 11 (third from the top). So, the Gray code of S_1 is 0101.

II.A.3 PSK

The system model is similar to that of an M -PAM constellation, the only difference being that the information rides on the phase of the symbol, rather than the amplitude. In addition to arranging the bits in a symbol w.r.t. their priorities, we need to place these symbols in a hierarchical way in the constellation to achieve unequal error protection. In Figure 4, the fictitious BPSK constellation denoted by the * symbol represents the highest priority (hereafter referred to as first level of hierarchy or sub-channel i_1). The angle θ_2 represents the second priority (second level of hierarchy or sub-channel i_2). Finally, the angle θ_3 represents the third priority (third level of hierarchy or sub-channel i_3). In other words, the constellation (A through H) can be visualized as a BPSK (S_1 and S_2) embedded into a QPSK (T_1 through T_4), which is further embedded into an 8-PSK (A through H). Gray coding is also done hierarchically. The same equations (1) can be used to Gray code a generalized hierarchical M -PSK constellation. For reader convenience, MATLAB programs automating the Gray code generation are available at [16].

II.B System Parameters

II.B.1 PAM

Distances

As we have described in section II-A.1, the distances we use evolve in a hierarchy. To simplify the notation in our proposed algorithm, we define the distance vector as $\mathbf{d} = [d_1 \ d_2 \ \cdots \ d_m]$ and the priority vector \mathbf{p} as

$$\mathbf{p} = [p_1 p_2 \cdots p_{m-1} p_m] = \begin{bmatrix} d_1 & d_2 & \cdots & d_{m-1} \\ d_m & d_m & \cdots & d_m \end{bmatrix} \mathbf{1}. \quad (2)$$

This vector controls the relative message priorities. The larger the ratio p_i / p_{i+1} the greater is the protection for the bit in position i than the bit in position $(i + 1)$.

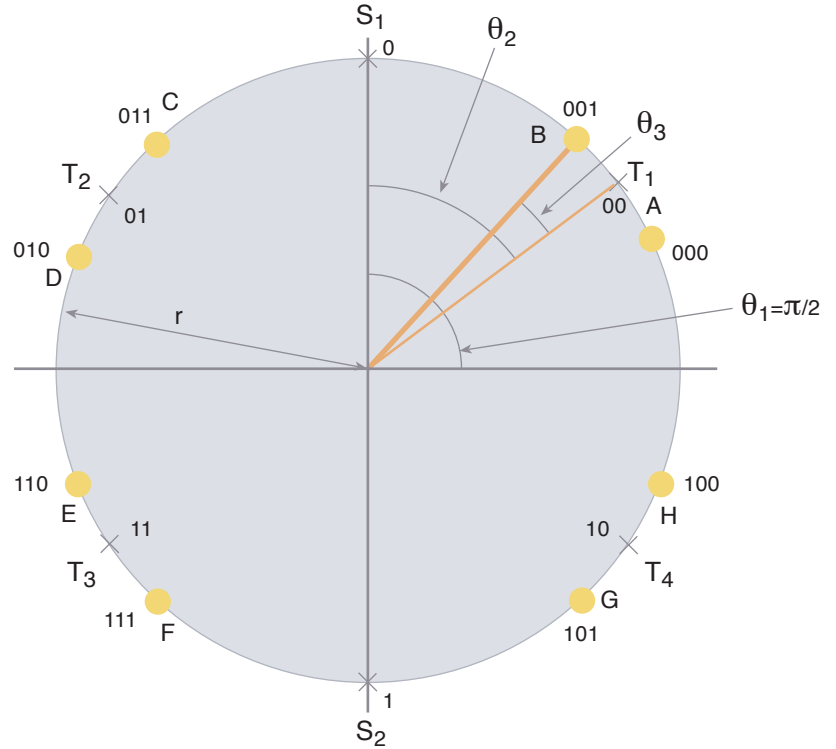


Figure 4 Generalized hierarchical 2/4/8-PSK constellations

Energies

The average symbol energy for the generalized M -PAM constellation can be computed as follows. If we were to denote every point in the constellation by its distance from the origin, then the coordinate will be of the form $\mathbf{p}\mathbf{a}^T$, where \mathbf{a} is a coordinate row vector (with elements $\pm d_m$). For example, the symbol S_1 in Figure 2 has vector $\mathbf{a} = [-1 \ 1 \ -1]d_3$. The energy of a point in the constellation with coordinate x can be written x^2 . By construction, we observe that for every point with coordinates $(d_1 + x)$, there is a point with coordinates $(d_1 - x)$. Also, we note that for every point with coordinates $(d_1 + d_2 + x)$, there exists a point with coordinates $(d_1 + d_2 - x)$, and this is true through the hierarchy from level 1 up to level m . Hence, because of this symmetry, we can see that the average energy E_s can be written as

$$E_s = (d_1^2 + d_2^2 + \cdots + d_m^2) = \mathbf{p}\mathbf{p}^T d_m^2. \quad (3)$$

II.B.2 QAM

Distances

As has been mentioned previously, square M -QAM constellations can be viewed as two \sqrt{M} -PAMs in quadrature. Therefore, to describe them, we need two distance vectors, \mathbf{d}^i and \mathbf{d}^q , where the superscripts “ i ” and “ q ” again refer to the I-phase and the Q-phase respectively.

In the case of square QAM, these two vectors are identically equal and are given by

$$\mathbf{d}^i = \mathbf{d}^q = [d_1 \ d_2 \ \dots \ d_m]. \quad (4)$$

Energies

Just as in the case of PAM, we can show for QAM that the average energy is given by:

$$E_s = 2 (d_1^2 + d_2^2 + \dots + d_m^2) = 2\mathbf{p}\mathbf{p}^T d_m^2. \quad (5)$$

II.B.3 PSK

As we have described in section II-A.3, a $2/4/\dots/M$ -PSK ($M = 2^m$) constellation can be described through a set of angles. To simplify the notation, we define an angle vector as

$$\boldsymbol{\theta} = [\theta_1 \ \theta_2 \ \dots \ \theta_m]_{1 \times m} \quad (6)$$

and the priority vector \mathbf{p} as

$$\mathbf{p} = [p_1 p_2 \ \dots \ p_{m-1} p_m] = \begin{bmatrix} \theta_1 & \theta_2 & \dots & \theta_{m-1} \\ \theta_m & \theta_m & \dots & \theta_m \end{bmatrix}_{1 \times m} \quad (7)$$

with $\theta_1 = \frac{\pi}{2}$. Note that these definitions are on

the same lines as for PAMs. As in the case of PAMs, the \mathbf{p} vector controls the relative message priorities. The larger the ratio p_i / p_{i+1} the greater is the protection for the bit in position i than the bit in position $(i + 1)$. Another important system parameter is the symbol energy for the generalized $2/4/\dots/M$ -PSK constellation, which is given by $E_s = r^2$, where r is the symbol amplitude.

II.B.4 Demodulator/Detector

PAM

The demodulator is based on Maximum Likelihood (ML) rule. The amplitude \hat{d} , of the incoming signal is recovered. This amplitude is then used for decoding the bit streams. We could interpret the decoder in two ways. We could use \hat{d} to identify the most likely transmitted ampli-

tude d , and get the m bits corresponding to this d and assign them to their respective bit streams (MSB to i_1, \dots , LSB to i_m). Equivalently, we

could use \hat{d} to decode individual bit streams as shown in Figure 5. We note from Figure 2 that for symbols $2/4/8$ -PAM constellation, the MSB is 0 in the right half plane, and 1 in the left half

plane. So, if the recovered amplitude, \hat{d} , is positive, then we can directly assign "0" to the MSB. Similarly, the next most significant bit (i_2)

is 1 if $|\hat{d}| \leq d_1$ and 0 else. So, if $\hat{d} \geq d_1$, we directly assign "0" to bit i_2 . More generally, for $2/4/\dots/M$ -PAM, the demodulator uses the following decision rules:

- For bit i_1 : If $\hat{d} \geq 0$, $i_1 = 0$; else $i_1 = 1$.
- For bit i_2 : If $|\hat{d}| \geq d_1$, $i_2 = 0$; else $i_2 = 1$.
- ...
- For bit i_m : The decision boundaries are given by vector \mathbf{B} , that can be constructed as shown by the MATLAB code in Table I.

This is nothing but ML decoding done on individual bits instead of symbols. Demodulation for QAMs can be done similarly.

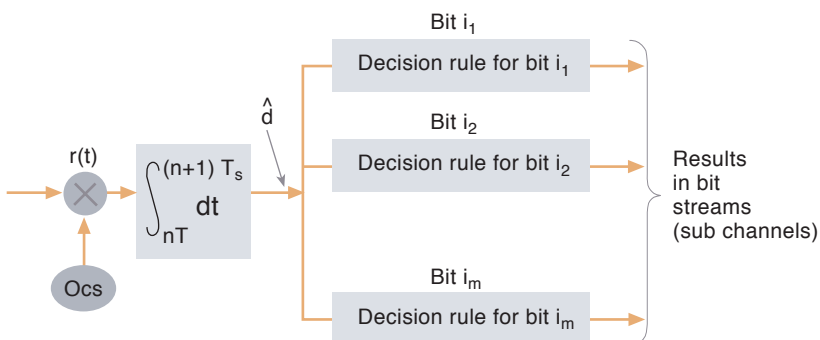
PSK

The demodulator is also based on the ML rule. The phase η , of the incoming signal is recovered. This phase is then used for decoding the bit streams. We could interpret the decoder in two ways. We could use η to identify the most likely transmitted phase θ , and get the m bits corresponding to this θ , and assign them to their respective bit streams (MSB to i_1, \dots , LSB to i_m). Equivalently, we could use θ to decode individual bit streams. We note from Figure 4 that for symbols $2/4/8$ -PSK constellation, the MSB is 0 in the upper half plane, and 1 in the lower half plane. So if the recovered phase, η , is such that $0 < \eta < \pi$, then we can directly assign "0" to the MSB. Similarly, the next most significant bit (i_2) is 1 in the left half plane, and 0 in the right half

plane. So if $-\frac{\pi}{2} < \eta < \frac{\pi}{2}$, we directly assign "0" to bit i_2 . More generally, for $2/4/\dots/M$ -PSK, the demodulator uses the following decision rules:

- For bit i_1 : If $0 < \eta < \pi$, $i_1 = 0$; else $i_1 = 1$.
- For bit i_2 : If $-\frac{\pi}{2} < \eta < \frac{\pi}{2}$, $i_2 = 0$; else $i_2 = 1$.
- ...
- For bit i_m : The decision boundaries are given by vector \mathbf{B} , that can be constructed similar to what is shown by the MATLAB code in Table I.

Figure 5 Hierarchical $2/4/\dots/M$ -PAM Demodulator



Once again, this is nothing but ML decoding done on individual bits instead of symbols.

III. Exact BER Computation

III.A Exact BER for 2/4/.../M-PAM Constellations

III.A.1 2/4-PAM Constellation

Consider the 2/4-PAM constellation as shown in Figure 1. The probability of error for the bit i_1 in

symbol 00 is given by $\frac{1}{2} \operatorname{erfc} \frac{d_1 + d_2}{\sqrt{N_0}}$, where

$\operatorname{erfc}(\cdot)$ function is the complimentary error function defined by

$$\operatorname{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^{\infty} \exp(-z^2) dz$$

and $\frac{N_0}{2}$ is the two-sided power spectral density

of the band pass AWGN. Similarly, the bit error probability for bit i_1 in symbol 01 is given by

$\frac{1}{2} \operatorname{erfc} \frac{d_1 - d_2}{\sqrt{N_0}}$. Hence, the average bit error

probability for bit i_1 , $P_b(4, (d_1, d_2); i_1)$, is given by

$$P_b(4, (d_1, d_2), i_1) = \frac{1}{4} \operatorname{erfc} \frac{d_1 - d_2}{\sqrt{N_0}} + \frac{1}{4} \operatorname{erfc} \frac{d_1 + d_2}{\sqrt{N_0}}. \quad (8)$$

Now, consider the second sub-channel i_2 . It can be shown that the average bit error probability for the sub-channel i_2 , $P_b(4, (d_1, d_2), i_2)$, is given by [17]

$$P_b(4, (d_1, d_2), i_2) = \frac{1}{4} \left[2 \operatorname{erfc} \frac{d_2}{\sqrt{N_0}} - \operatorname{erfc} \frac{2d_1 + d_2}{\sqrt{N_0}} + \operatorname{erfc} \frac{2d_1 - d_2}{\sqrt{N_0}} \right]. \quad (9)$$

III.A.2 Generalized 2/4/.../M-PAM Constellation

We propose a recursive algorithm for the exact BER computation of generalized 2/4/.../M-PAM constellations. We use the generalized 4-PAM constellation as the root for this algorithm. We have already developed the expressions for exact BER of bits i_1 and i_2 in the case of 4-PAM ($m = 2$). To compute the BER for bits i_k , $P_b(M, \mathbf{d}, i_k)$, ($k = 1, 2, \dots, m$), where bit i_m represents the LSB and $m > 2$. The pseudo-code of the algorithm was shown to be given by [17]:

- If $k < m$

$$P_b(M, \mathbf{d}, i_k) = \frac{1}{2} \left[P_b\left(\frac{M}{2}, \mathbf{d}_+, i_k\right) + P_b\left(\frac{M}{2}, \mathbf{d}_-, i_k\right) \right], \quad (10)$$

Generation of vector \mathbf{B}

```

DecisionBoundVector=d(1 : m - 1)
end
B=zeros(2m-1, 1)
for i = 2m-2 : 2m-1 - 1
    index=i
    mult=zeros(1, m - 1)
    for j = m - 1 : -1 : 1
        mult(1, j)=2*mod(index,2)-1
    end
    index=floor(index/2)
end
B(i + 1,1)=mult*DecisionBoundVector
end
for i = 2m-2 : -1 : 1
    B(i, 1) = -B(2m-1 - i + 1, 1)
end

```

Table I MATLAB pseudo-code for the generation of vector \mathbf{B}

where

$$\mathbf{d}_+ = [d_1 \ d_2 \ \dots \ d_{m-2} \ d_{m-1} + d_m]_{1 \times (m-1)} \quad (11)$$

and

$$\mathbf{d}_- = [d_1 \ d_2 \ \dots \ d_{m-2} \ d_{m-1} - d_m]_{1 \times (m-1)} \quad (12)$$

are $(m - 1)$ dimensional row vectors (\mathbf{d} is an m dimensional row vector). As we proceed through the recursion, the constellation size goes on decreasing until we either reach a stage wherein the vectors \mathbf{d}_+ and \mathbf{d}_- are of length 2 (in the case of bits i_1 and i_2), or bit i_k (for $k > 2$) becomes the LSB. In the former case we use the result from 4-PAM (given in Sect. III-A.1), and come out of the recursion. In the latter case ($k = m$), we do the following.

- If $k = m$ or in the event that bit i_k becomes the LSB at some stage through the recursion, we need an LSB recursion algorithm. It was shown in [17] that the recursion for the LSB is given by

$$P_b(M, \mathbf{d}, i_m) = \frac{1}{2^m} (P_0 + P_1), \quad (13)$$

where

$$P_0 = \sum_{i=1}^{2^{m-1}} \sum_{j=1}^{2^{m-1}} \frac{1}{2} \left[(-1)^{j+1} \operatorname{erfc} \left(\frac{\mathbf{B}(j) - \mathbf{d}_0(i)}{\sqrt{N_0}} \right) \right] \quad (14)$$

and

$$P_1 = \sum_{i=1}^{2^{m-1}} \left(1 + \sum_{j=1}^{2^{m-1}} \frac{1}{2} \left[(-1)^j \operatorname{erfc} \left(\frac{\mathbf{B}(j) - \mathbf{d}_1(i)}{\sqrt{N_0}} \right) \right] \right), \quad (15)$$

Table II MATLAB pseudo-code for generation of vectors \mathbf{d}_0 and \mathbf{d}_1

Generation of vectors \mathbf{d}_0 and \mathbf{d}_1

```

 $\mathbf{d}_{sym} = \text{zeros}(2^m, 1)$ 
 $\mathbf{d}_0 = \text{zeros}(2^{m-1}, 1)$ 
 $\mathbf{d}_1 = \text{zeros}(2^{m-1}, 1)$ 
for  $i = 2^{m-1} : 2^m - 1$ 
    index =  $i$ 
    mult = zeros(1,  $m - 1$ )
    for  $j = m - 1 : -1 : 1$ 
        mult(1,  $j$ ) =  $2 * \text{mod}(\text{index}, 2) - 1$ 
    end
    index = floor(index/2)
end
 $\mathbf{d}_{sym}(i + 1, 1) = \text{mult} * \mathbf{d}$ 
end
for  $i = 1 : 2^{m-1}$ 
     $\mathbf{d}_{sym}(i, 1) = -\mathbf{d}(2^m - i + 1, 1)$ 
end
 $\mathbf{d}_0(1, 1) = \mathbf{d}_{sym}(1, 1)$ ;
 $x = 2; y = 1$ ;
while ( $x < 2^m - 1$ )
     $\mathbf{d}_1(y, 1) = \mathbf{d}_{sym}(x, 1)$ 
     $\mathbf{d}_1(y + 1, 1) = \mathbf{d}_{sym}(x + 1, 1)$ 
     $\mathbf{d}_0(y + 1, 1) = \mathbf{d}_{sym}(x + 2, 1)$ 
     $\mathbf{d}_0(y + 2, 1) = \mathbf{d}_{sym}(x + 3, 1)$ 
     $x = x + 4$ 
     $y = y + 2$ 
end
 $\mathbf{d}_1(y, 1) = \mathbf{d}_{sym}(x, 1)$ 
 $\mathbf{d}_1(y + 1, 1) = \mathbf{d}_{sym}(x + 1, 1)$ 
 $\mathbf{d}_0(y + 1, 1) = \mathbf{d}_{sym}(x + 2, 1)$ 

```

where

- \mathbf{B} is the vector of decision boundary positions for the LSB, w.r.t. the origin,
- \mathbf{d}_0 is the vector of positions of constellation points whose LSB is 0, w.r.t. the origin, and
- \mathbf{d}_1 is the vector of positions of constellation points whose LSB is 1, w.r.t. the origin.

The generation of these vectors is shown in Tables I and II. For the reader's convenience, our MATLAB computer programs for PAM (along with a readme file) are available at [15] to allow one to immediately compare the performance of various generalized hierarchical constellations.

III.B Extension to Generalized Hierarchical M -QAM Constellations

As mentioned before, QAM constellations can be viewed as 2 PAMs in quadrature. This fact helps to deduce the BER expressions [17] for QAMs in terms of those of the 2 PAM constellations. For instance, square QAM constellations use 2 bits for every level of priority ($M = 2^{2m}$). Figure 6 shows a 4/16/64-QAM constellation.

In the case where every level of priority is not restricted to 2 bits per channel access, many other QAM constellations can arise. The recursive algorithm developed here for the exact BER computation can be readily adapted to treat all these cases. As an example of the applicability of the proposed algorithm, we will consider a family of rectangular M -QAM (i.e. $M = 2^{2m+1}$) constellations with $m + 1$ incoming streams of data. For this considered family of constellations the highest priority level is assigned 1 bit in the in-phase whereas all the other levels are assigned two bits in similar fashion as the square QAM case described above. This can be viewed as a 2/8/32.../ M -QAM constellation. As an illustration of this family of constellations Figure 7 shows a generalized 32-QAM constellation.

III.C Exact BER for 2/4/.../ M -PSK

III.C.1 2/4-PSK Constellation

Consider the 2/4-PSK constellation shown in Figure 8. For the MSB (bit i_1), it can be shown that

$$P_b(4, (\theta_1, \theta_2), i_1) = \frac{1}{2} \text{erfc}(\sqrt{\gamma} \sin \theta_2), \quad (16)$$

where $\gamma = \frac{r^2}{N_0}$ is the symbol carrier to noise

ratio (CNR). Now, consider the LSB (second sub-channel i_2). It can easily be shown that [18]

$$P_b(4, (\theta_1, \theta_2), i_2) = \frac{1}{2} \text{erfc}(\sqrt{\gamma} \cos \theta_2). \quad (17)$$

III.C.2 Generalized 2/4/.../ M -PSK Constellations

We notice that the decision boundaries for bits i_1, i_2, \dots, i_m in terms of $\theta_1, \theta_2, \dots, \theta_m$ are the same in both 2/4/.../ M -PSK and 2/4/.../ $M/2M$ -PSK. So, the protection angles for these bits in the latter case differ from those in the former, by $\pm \theta_{m+1}$. Similar to the PAM case, we propose a recursive algorithm for which we use the generalized 2/4-PSK constellation as a root. We have already developed the expressions for exact BER of bits i_1 and i_2 in the case of 2/4-PSK ($m = 2$). To compute the BER for bits i_k , $P_b(M, \theta, i_k)$, ($k = 1, 2, \dots, m$), where bit i_m represents the LSB, and

$m > 2$, we use a recursive algorithm, whose pseudo code is given as follows:

- If $k < m$

$$P_b(M, \theta, i_k) = \frac{1}{2} \left[P_b\left(\frac{M}{2}, \theta_+, i_k\right) + P_b\left(\frac{M}{2}, \theta_-, i_k\right) \right], \quad (18)$$

where

$$\theta_+ = [\theta_1 \theta_2 \cdots \theta_{m-2} \theta_{m-1} + \theta_m]_{1 \times (m-1)} \quad (19)$$

and

$$\theta_- = [\theta_1 \theta_2 \cdots \theta_{m-2} \theta_{m-1} - \theta_m]_{1 \times (m-1)} \quad (20)$$

are $(m - 1)$ dimensional row vectors (note that θ is an m dimensional row vector). As we proceed through the recursion, the constellation size goes on decreasing until we either reach a stage wherein the vectors θ_+ and θ_- are of length 2 (this is the case with bits i_1 and i_2) or the bit i_k (for $k > 2$) becomes the LSB. In the former case we use the result from 2/4-PSK (given in Section III-C.1) and come out of the recursion. In the latter case, we use the following closed form expression.

- If $k = m$ or in the event that bit i_k becomes the LSB at some stage through the recursion, the BER of the LSB, $P_b(M, \theta, i_m)$, can be written in closed form [18] as

$$P_b(M, \theta, i_m) = \frac{1}{2^m} (P_0 + P_1), \quad (21)$$

where

$$P_0 = \sum_{i=1}^{2^{m-1}} \sum_{j=1}^{2^{m-1}} (-1)^j F(\mathbf{B}(j) - \phi_0(i)), \quad (22)$$

and

$$P_1 = \sum_{i=1}^{2^{m-1}} \sum_{j=1}^{2^{m-1}} (-1)^{j+1} F(\mathbf{B}(j) - \phi_1(i)), \quad (23)$$

where the F-function is defined in [19], [20] as follows:

$$F(\psi) = -\frac{\text{sgn}(\psi)}{2\pi} \int_0^{\pi-|\psi|} \exp\left[-\gamma \frac{\sin^2 \psi}{\sin^2 \theta}\right] d\theta, \quad (24)$$

$-\pi < \psi < \pi,$

where $\text{sgn}(\cdot)$ is the sign function. Please refer to [19], [20] for the properties of this function.

In (22) and (23), the vectors \mathbf{B} , ϕ_0 and ϕ_1 are defined as follows.

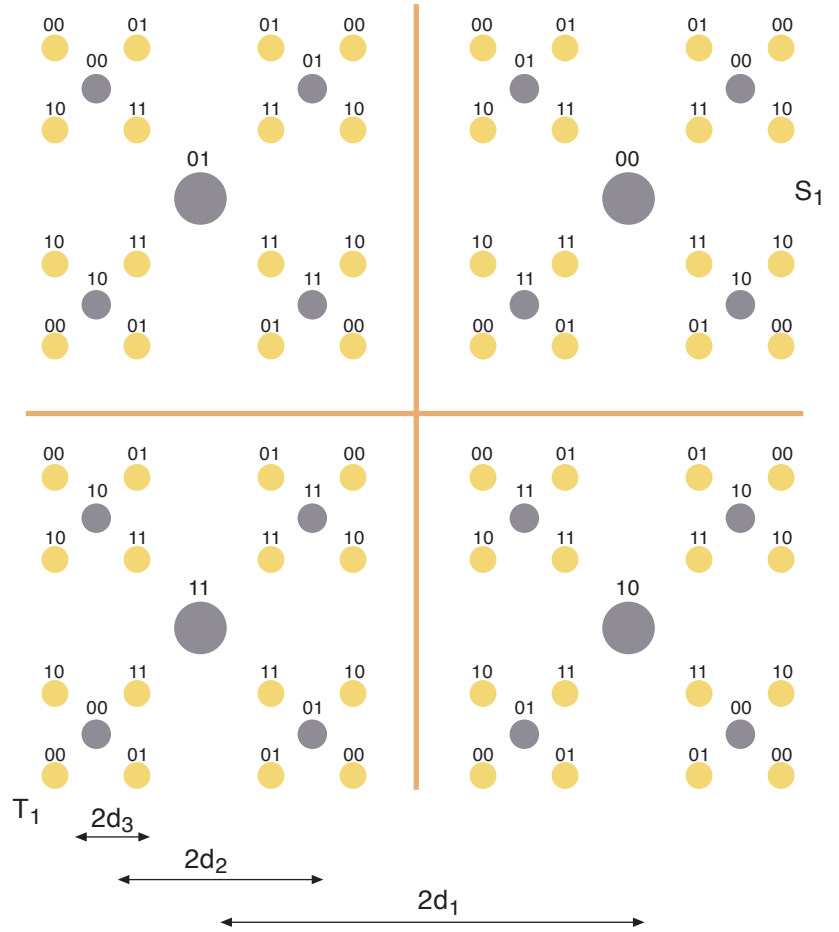


Figure 6 Hierarchical 64-QAM constellation. Gray coding is done hierarchically. For instance, the Gray code for symbol S_1 is 001010, while that for symbol T_1 is 110000

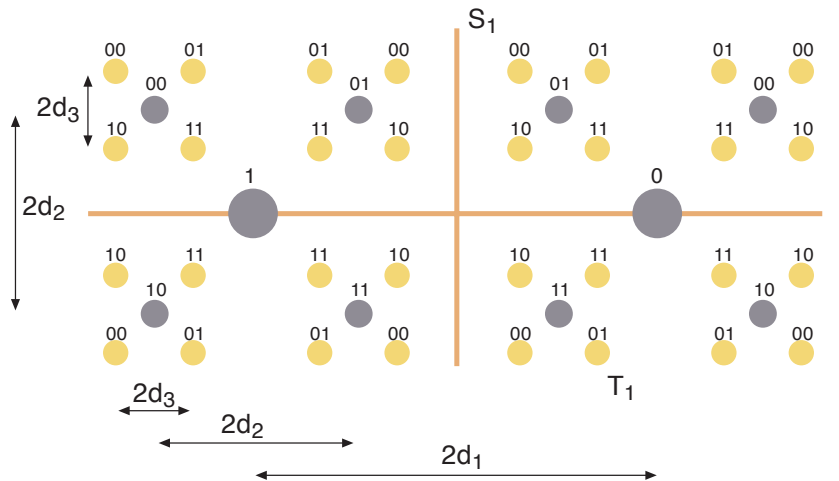


Figure 7 Hierarchical 32-QAM constellation. The Gray coding is done hierarchically. For instance, the Gray code for symbol S_1 is 00100, while that for symbol T_1 is 01101

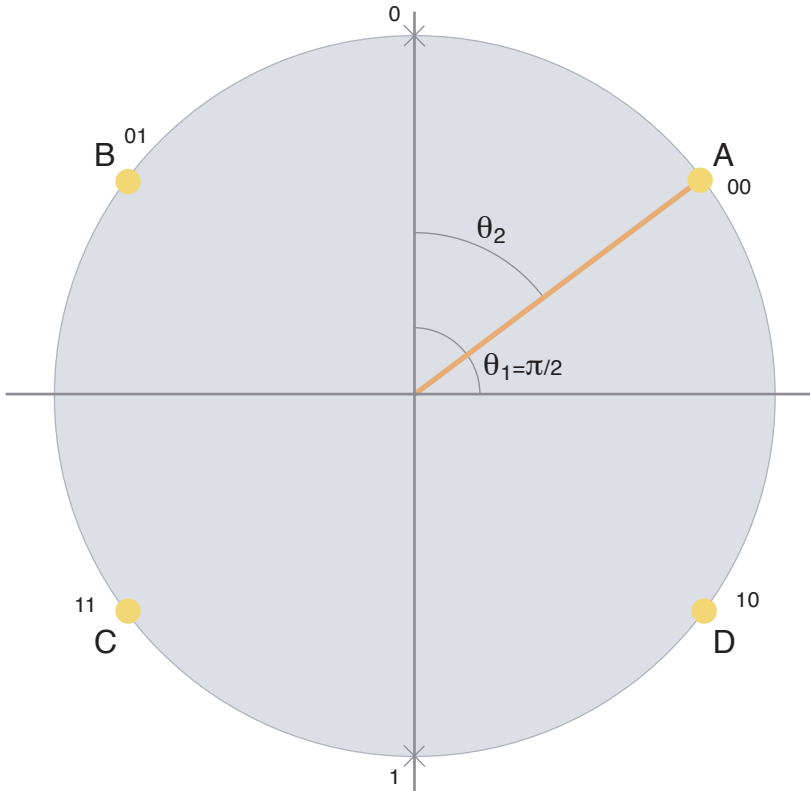


Figure 8 2/4-PSK constellation

- The elements of \mathbf{B} represent the angular positions of the decision boundaries of the LSB w.r.t. the reference axis.
- The elements of ϕ_0 represent the angular positions of those symbols in the constellation ($2^m - \text{PSK}$), whose LSB is '0', w.r.t. the reference axis.
- The elements of ϕ_1 represent the angular positions of those symbols in the constellation ($2^m - \text{PSK}$), whose LSB is '1', w.r.t. the reference axis.

As an example, see Figure 9 for the generation of vectors \mathbf{a} , ϕ_0 and ϕ_1 in the case of 2/4/8-PSK constellation. More generally, these vectors can be easily generated for any 2/4/.../M-PSK. Note that vector \mathbf{B} is similar to vector \mathbf{B} in the case of PAMs, and it can be generated using the same pseudo code given in Table I, but by replacing vector \mathbf{d} with θ . Vectors ϕ_0 and ϕ_1 are similar to the vectors \mathbf{d}_0 and \mathbf{d}_1 respectively. They can be generated using the same pseudo-code given in Table II, but by replacing vectors \mathbf{d} , \mathbf{d}_0 and \mathbf{d}_1 with θ , ϕ_0 and ϕ_1 respectively. For reader convenience, all our MATLAB programs for PSK (along with a readme file) are available at [16].

IV. Special Cases

IV.A Uniform M-PAM

Standard uniform M -PAM constellations can be viewed as a special case of the generalized M -PAMs when $\mathbf{p} = [2^{m-1} 2^{m-2} \dots 4 2 1]$. The resulting BER for bit i_k , $P_b(M, \mathbf{p}, i_k)$, obtained by using the recursive algorithm, can be shown to be in agreement with the explicit closed form expression recently derived in [13] and given by

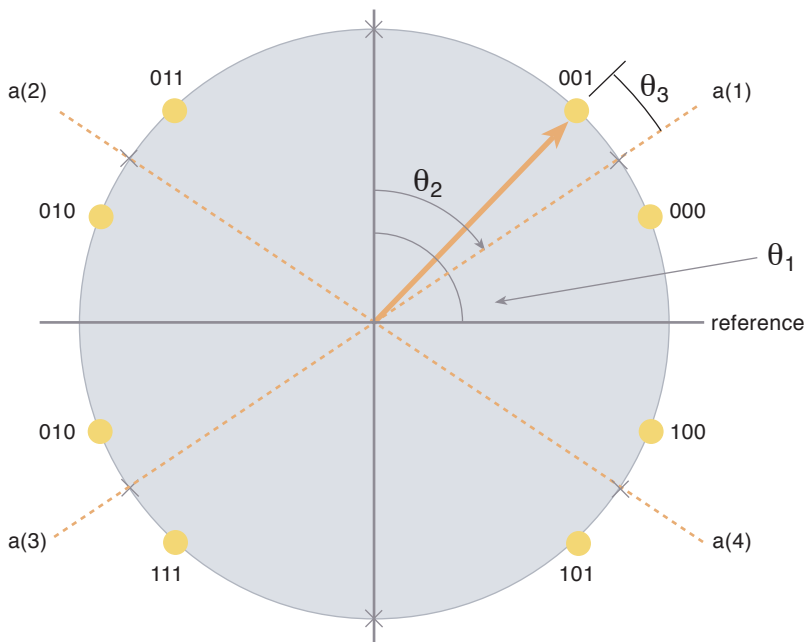
$$P_b(M, \mathbf{p}, i_k) = \frac{1}{M} \sum_{j=0}^{(1-2^{-k})M-1} (-1)^{\lfloor \frac{j2^{k-1}}{M} \rfloor} \left[2^{k-1} - \left\lfloor \frac{j2^{k-1}}{M} + \frac{1}{2} \right\rfloor \right] \operatorname{erfc} \left[(2j+1) \sqrt{\frac{3\gamma \log_2 M}{(M^2-1)}} \right], \quad (25)$$

where $\lfloor \cdot \rfloor$ is the floor function. The average BER $P_b(M, \mathbf{p})$ is given by

$$P_b(M, \mathbf{p}) = \frac{1}{m} \sum_{k=1}^m P_b(M, \mathbf{p}, i_k). \quad (26)$$

IV.B Uniform Square QAM

M -QAM can be viewed as a special case of the generalized M -QAMs when $\mathbf{p} = [2^{m-1} 2^{m-2} \dots 4 2 1]$. Please refer to Figure 6 for a 4/16/64-QAM example. This can be used when all the incoming bit streams have nearly the same priority.



$$\text{Vector } \mathbf{a} = \left[\frac{\pi}{2} - \theta_2 \quad \frac{\pi}{2} + \theta_2 \quad -\frac{\pi}{2} - \theta_2 \quad -\frac{\pi}{2} + \theta_2 \right]$$

$$\text{Vector } \Phi_0 = \left[\frac{\pi}{2} - \theta_2 - \theta_3 \quad \frac{\pi}{2} + \theta_2 + \theta_3 \quad -\frac{\pi}{2} - \theta_2 - \theta_3 \quad -\frac{\pi}{2} + \theta_2 + \theta_3 \right]$$

$$\text{Vector } \Phi_1 = \left[\frac{\pi}{2} - \theta_2 - \theta_3 \quad \frac{\pi}{2} + \theta_2 - \theta_3 \quad -\frac{\pi}{2} - \theta_2 - \theta_3 \quad -\frac{\pi}{2} + \theta_2 - \theta_3 \right]$$

Figure 9 Vectors \mathbf{a} , ϕ_0 and ϕ_1 for a 2/4/8-PSK constellation

By symmetry, the in-phase and quadrature phase bits (i_k and q_k) have the same BER. The average BER for both the in-phase and quadrature phase bits is obtained by averaging the BER of all in-phase bits i_k and quadrature bits q_k

($k = 1, \dots, \frac{1}{2} \log_2 M$) yielding

$$P_b^s(M, \mathbf{d}) = \frac{1}{2m} \left[\sum_{k=1}^m P_b(\sqrt{M}, \mathbf{d}, i_k) + \sum_{k=1}^m P_b(\sqrt{M}, \mathbf{d}, q_k) \right] = \frac{1}{m} \sum_{k=1}^m P_b(\sqrt{M}, \mathbf{d}, i_k), \quad (27)$$

where $P_b(\sqrt{M}, \mathbf{d}, i_k)$ is the BER of the bit i_k in a \sqrt{M} -PAM constellation.

IV.C Uniform PSK

Uniform M-PSK can be viewed as a special case of the generalized 2/4/.../M-PSKs when

$\theta = \left[\frac{\pi}{2} \frac{\pi}{4} \frac{\pi}{8} \dots \frac{\pi}{M} \right]$, (i.e. when $\mathbf{p} = [2^{m-1} 2^{m-2} 2^{m-3} \dots 4 2 1]$). The average BER is obtained by averaging the BER of all the bits, i_k ($k = 1, 2, \dots, \log_2 M$) yielding

$$P_b(M) = \frac{1}{\log_2 M} \left[\sum_{k=1}^{\log_2 M} P_b(M, \theta, i_k) \right] \quad (28)$$

where $P_b(M, \theta, i_k)$ is the BER of the bit i_k in the constellation, and which can be evaluated using (18) and (21). Using these equations in (28), it can also be shown that the exact BER of uniform M-PSK can be written in terms of the Hamming weights of the M symbols, W_k^M , ($k = 1, 2, \dots, M$) as

$$P_b(M) = \frac{1}{\log_2 M} \sum_{k=1}^M W_k^M \left[F \left(\frac{(2k-1)\pi}{M} \right) - F \left(\frac{(2k-3)\pi}{M} \right) \right]$$

in agreement with [21], [22] and [23, Section 4.1]. Note that for constellations Gray coded in the Karnaugh map style (as shown in section II-A.3), these Hamming weights W_k^M can be calculated in terms of the m bits ($m = \log_2 M$) in the Gray codes of the symbols, g_k , (corresponding to the k -th symbol) as

$$W_k^M = \sum_{i=1}^{\log_2 M} g_{i,k}$$

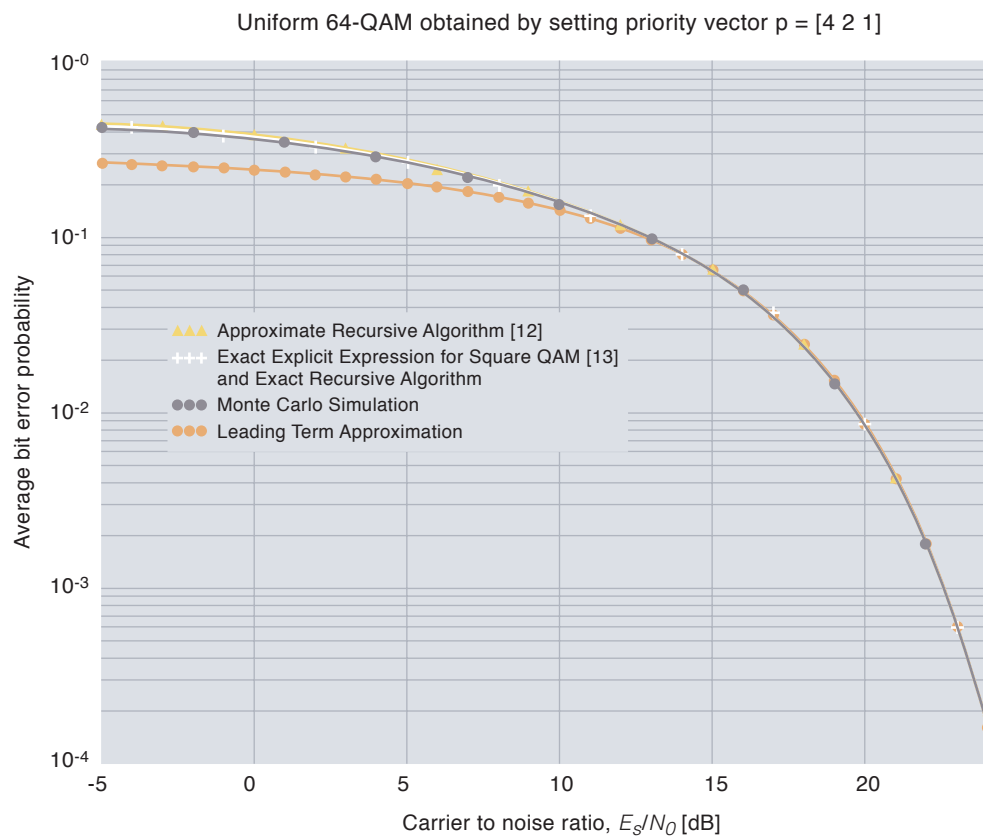


Figure 10 Comparison of the various methods (Yang and Hanzo [12], Cho and Yoon [13] and the proposed recursive algorithm) for the computation of the BER for the uniform 64-QAM case with $\mathbf{p} = [4 2 1]$

V. Numerical Examples

Since the number of different cases covered by our analysis is quite large, we just present in this section some numerical examples that demonstrate the usefulness of the proposed analytical

tools for a variety of scenarios of practical interest. We further note that the analytical expressions derived in the previous sections, and the corresponding numerical results presented in this section have been verified extensively by Monte

Figure 11 Performance of 4/16/64-QAM. Note that the MSB performs much better than the LSB

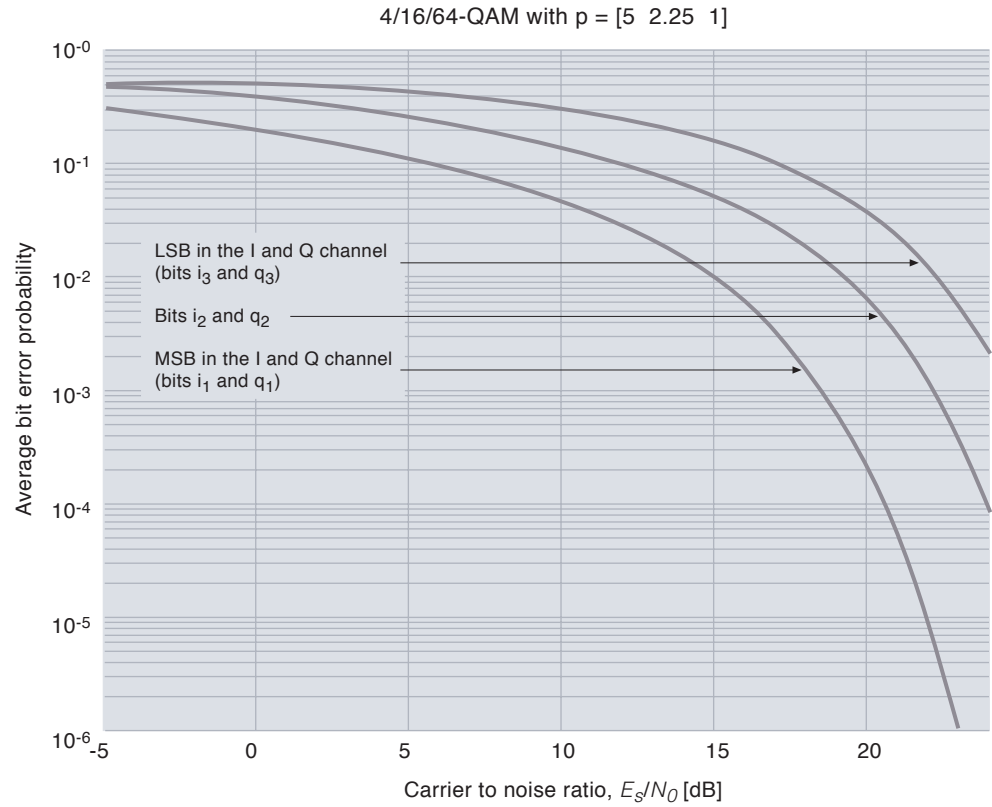
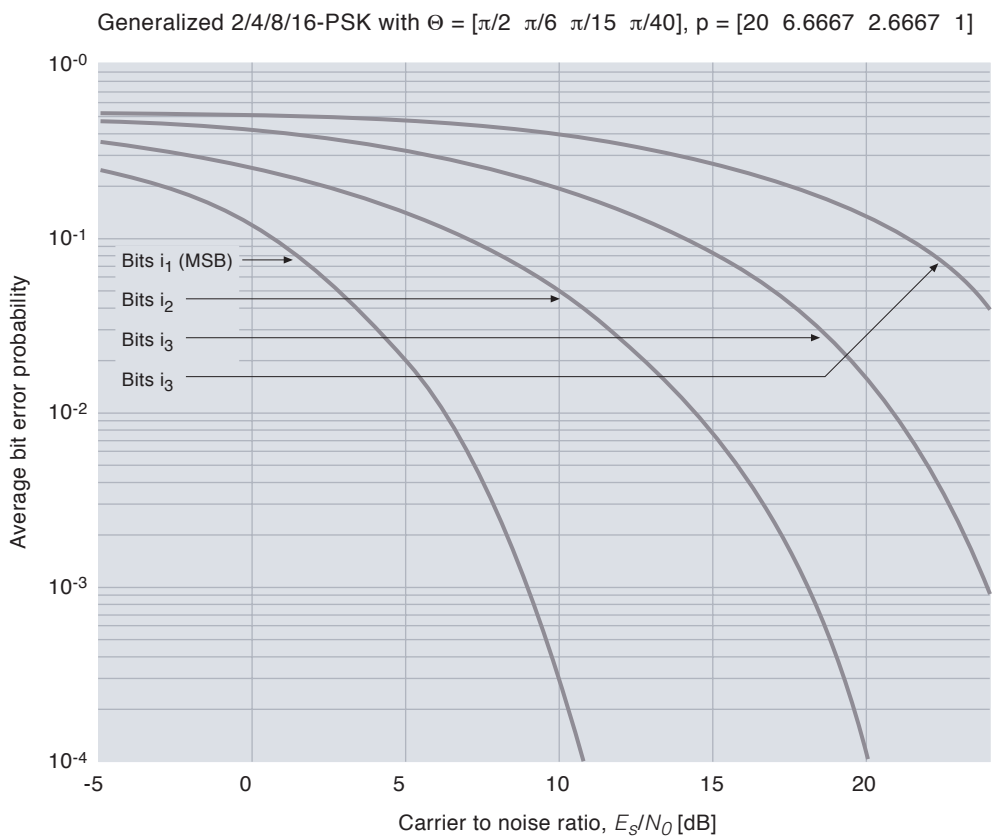


Figure 12 Performance of 2/4/8/16-PSK. Note that the MSB performs nearly 20 dB better than the LSB at $BER \leq 10^{-2}$



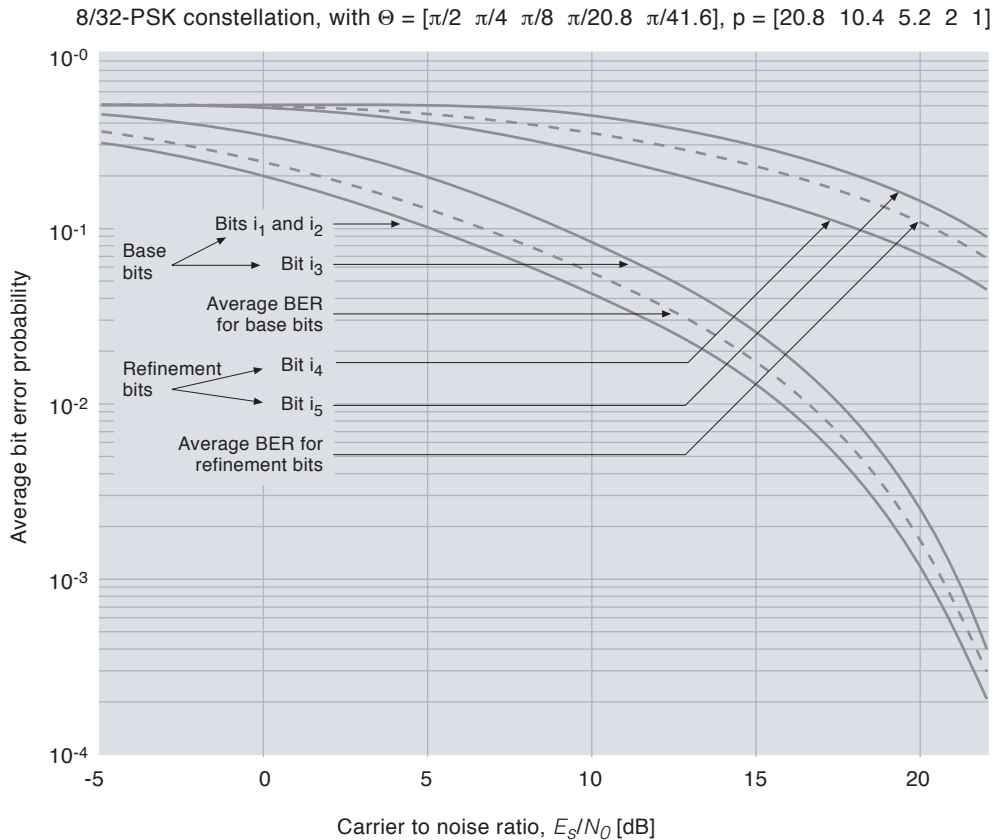


Figure 13 Performance of 2/4/8/16/32-PSK. The 3 most significant bits are the base bits and the remaining are refinement bits. Note that on average, the base bits perform 10–15 dB better than the refinement bits at $\text{BER} \leq 10^{-2}$

Carlo simulations for the various constellations under consideration. Thus, the reader can be totally confident in the correctness of the newly derived recursive expressions and the accuracy of the numerical results illustrated below.

Figure 10 deals with the uniform 64-QAM case and shows the perfect match between the proposed exact recursive algorithm and the exact generic expression obtained by Yoon et al. [13] as well as the Monte-Carlo simulations. The approximate recursive expression obtained in [12] comes very close to the exact result, whereas the leading term approximation¹⁾ gives an optimistic result at low CNR. Figure 11 shows the performance of 4/16/64-QAM constellation. Note that for the priority vector $\mathbf{p} = [5 \ 2.25 \ 1]$, the protection of the MSB gets a higher protection than the LSB. Figure 12 shows the similar nature of performance of 2/4/8/16-PSK. Figure 13 shows the performance of a 2/4/8/16/32-PSK constellation, where the \mathbf{p} vector is given by $\mathbf{p} = [\beta^4 \ \beta^3 \ \beta^2 \ 2 \ 1]$, with $\beta = 1.3 \geq 1$. This is a situation wherein the 3 most significant bits are protected better than the 2 least significant bits. Also, note that the bits i_1, i_2, i_3 form a uniform 8-PSK, while the bits i_4, i_5 form a uniform 4-PSK. So, this is a constellation wherein an 8-PSK, embedded in a 32-PSK is protected by

the factor β . These protected bits (i_1, i_2, i_3) are referred to as base bits, and the others (i_4, i_5) are referred to as the refinement bits. As β is increased, the 2 sets of curves move further apart.

VI. Potential Applications

VI.A Voice and Data Applications

Four years ago, a new scheme for simultaneous data and voice transmission over fading channels was proposed [24]. It provided high average spectral efficiency for data transmission, while meeting stringent delay requirements for voice transmission. This modem always transmits voice packets, in the form of a BPSK signal on the Q-channel. When fading is not severe, the modem assigns a significant transmit power to data transmission. This is done by using a large PAM constellation (on the I-channel). As the channel fading gets severe, the modem reduces the size of this constellation, thereby giving more power to voice transmission to meet the stringent delay constraints. We are currently looking into designing an alternative adaptive scheme wherein hierarchical QAM or PSK constellations are used instead, in order to improve the spectral efficiency for simultaneous multimedia transmission.

¹⁾ By leading term, we mean just the dominant $\text{erfc}(\cdot)$ term.

VI.B Downlink Multiplexing/ Multicasting

As we have seen in the previous section, by virtue of Gray coding, we protect the MSB much more than the LSB. Could this property along with hierarchical channel coding be used for multicasting information from base station to mobile units? Information for the user who suffers the least fading could be transmitted on the LSBs of the signal, while that for the user suffering the worst fading could be transmitted on the MSBs of the same signal. The users receive the symbol and look only for the bits in particular positions within the symbol. We are now focusing on the development of this application and the comparison of its spectral efficiency performance with that of existing multicast strategies.

VII. Conclusion

We have argued that the recursive way of Gray coding a PAM/PSK constellation ensures the existence of a recursive algorithm for the computation of the BER. As a result, we have obtained exact and generic expressions (in M) for the BER of the generalized hierarchical M -PAM constellations over AWGN channels. These results can be extended easily to fading channels [17], [18]. We have also shown that these expressions can be extended to generalized hierarchical M -QAM constellations. Using the same argument, we have derived similar BER expressions for hierarchical PSK constellations. Because of their generic nature, these new expressions readily allow numerical evaluation for various cases of practical interest. In particular numerical examples have shown that the leading-term approximation gives significantly optimistic BER values at low CNR but is quite accurate in the high CNR region. Finally, we have suggested possible applications, voice and data integration and downlink multiplexing, to name a few.

References

- 1 Cover, T. Broadcast channels. *IEEE Trans. on Inform. Theory*, IT-18, 2–14, 1972.
- 2 Sundberg, C E W, Wong, W C, Steele, R. Logarithmic PCM weighted QAM transmission over Gaussian and Rayleigh fading channels. *IEE Proc.*, 134, 557–570, 1987.
- 3 Ramchandran, K et al. Multiresolution broadcast for digital HDTV using joint source/channel coding. *IEEE Journal of Selected Areas in Communications*, 11, 1993.
- 4 Wei, L-F. Coded modulation with unequal error protection. *IEEE Trans. Commun.*, COM-41, 1439–1449, 1993.
- 5 Morimoto, M et al. A study on power assignment of hierarchical modulation schemes for digital broadcasting. *IEICE Trans. Commun.*, E77-B, 1994.
- 6 Morimoto, M, Okada, M, Komaki, S. A hierarchical image transmission system in a fading channel. In: *Proc. IEEE Int. Conf. Univ. Personal Comm. (ICUPC'95)*, 769–772, 1995.
- 7 Pursley, M B, Shea, J M. Nonuniform phase-shift-key modulation for multimedia multicast transmission in mobile wireless networks. *IEEE Journal of Selected Areas in Communications*, SAC-17, 774–783, 1999.
- 8 Pursley, M B, Shea, J M. Adaptive nonuniform phase-shift-key modulation for multimedia traffic in wireless networks. *IEEE Journal of Selected Areas in Communication*, SAC-00, 1394–1407, 2000.
- 9 DVB-T standard: ETS 300 744. *Digital Broadcasting Systems for Television, Sound and Data Services: Framing Structure, Channel Coding and Modulation for Digital Terrestrial Television*. (ETSI Draft, 1.2.1, EN300 744, 1999-1.)
- 10 Fitz, M O, Seymour, J P. On the bit error probability of QAM modulation. *International Journal of Wireless Information Networks*, 1 (2), 131–139, 1994.
- 11 Lu, J et al. M-PSK and M-QAM ber computation using signal-space concepts. *IEEE Trans. Commun.*, COM-47, 181–184, 1999.
- 12 Yang, L L, Hanzo, L. A recursive algorithm for the error probability evaluation of M-QAM. *IEEE Commun. Letters*, 4, 304–306, 2000.
- 13 Yoon, D, Cho, K, Lee, J. Bit error probability of M-ary quadrature amplitude modulation. In: *Proc. IEEE Veh. Technol. Conf. (VTC'2000-Fall)*, Boston, Massachusetts, 2422–2427, September 2000.
- 14 Vitthaladevuni, P K, Alouini, M-S. BER computation of generalized hierarchical 4/M-QAM constellations. In: *Proc. IEEE International Symposium on Personal, Indoor and Mobile Radio Commun. Conf. (PIMRC'2001)*, San Diego, California, 1, 85–89, September 2001. (Journal version in *IEEE Trans. on Broadcasting*, 47 (3), 228–239, 2001.)
- 15 Vitthaladevuni, P K, Alouini, M-S. *MATLAB programs for design and BER computation*

- of generalized hierarchical M -QAM constellations. Available at: <http://www.ece.umn.edu/users/pavan/Generalized-Hierarchical-Qam.html>
- 16 Vitthaladevuni, P K, Alouini, M-S. *MATLAB programs for design and BER computation of generalized hierarchical M -PSK constellations*. Available at: <http://www.ece.umn.edu/users/pavan/Generalized-Hierarchical-PSK.html>
 - 17 Vitthaladevuni, P K, Alouini, M-S. BER computation of generalized QAM constellations. In: *Proc. IEEE Global Commun. Conf. (GLOBECOM'2001)*, San Antonio, Texas, 1, 632–636, 2001. (Journal version submitted to *IEEE Trans. on Information Theory*.)
 - 18 Vitthaladevuni, P K, Alouini, M-S. BER computation of generalized hierarchical PSK constellations. In: *Proc. IEEE International Conference on Communications (ICC'2002)*, New York, April 2002. (Journal version submitted to *IEEE Trans. Commun.*)
 - 19 Pawula, R F, Rice, S O, Roberts, J H. Distribution of the phase angle between two vectors perturbed by Gaussian noise. *IEEE Trans. Commun.*, COM-30, 1828–1841, 1982.
 - 20 Pawula, R F. Distribution of the phase angle between two vectors perturbed by Gaussian noise II. *IEEE Trans. Veh. Technol.*, 50, 576–583, 2001.
 - 21 Lee, P J. Computation of the bit error rate of coherent M -ary PSK with Gray code bit mapping. *IEEE Trans. Commun.*, COM-34, 488–491, 1986.
 - 22 Tellambura, C, Mueller, A J, Bhargava, V K. Analysis of M -ary phase-shift keying with diversity reception for land-mobile satellite channels. *IEEE Trans. Veh. Technol.*, VT-46, 910–922, 1997.
 - 23 Simon, M K, Hinedi, S M, Lindsey, W C. *Digital Communication Techniques – Signal Design and Detection*. Englewood Cliffs, NJ, Prentice Hall, 1995.
 - 24 Alouini, M-S, Tang, X, Goldsmith, A. An adaptive modulation scheme for simultaneous voice and data transmission over fading channels. *IEEE Journal of Selected Areas in Communication*, SAC-17, 837–850, May 1999. (See also *Proc. of the IEEE Vehicular Technology Conference (VTC'98)*, 939-943, Ottawa, Ontario, Canada, May 1998.)

Performance Analysis of Adaptive Coded Modulation with Antenna Diversity and Feedback Delay

KJELL J. HOLE, HENRIK HOLM AND GEIR E. ØIEN



Kjell Jørgen Hole (41) received his BSc, MSc and PhD degrees in computer science from the University of Bergen in 1984, 1987 and 1991, respectively. He is currently Senior Research Scientist at the Department of Telecommunications at the Norwegian University of Science and Technology (NTNU) in Trondheim. His research interests are in the areas of coding theory and wireless communications.

Kjell.Hole@ii.uib.no



Henrik Holm (30) received his Siv.Ing. and PhD degrees in electrical engineering from the Norwegian University of Science and Technology (NTNU) in 1997 and 2002, respectively. He is currently a post doctoral researcher at the NTNU and a guest researcher at the University of Minnesota. His research interests include statistical modeling and robust transmission on wireless channels.

henrik@tele.ntnu.no

A general adaptive coding scheme for spectrally efficient transmission on flat fading channels was introduced by the authors in an earlier paper [2]. An instance of the coding scheme utilizes a set of multi-dimensional trellis codes designed for additive white Gaussian noise channels of different qualities. A feedback channel between the decoder and encoder makes it possible for the encoder to switch adaptively between these codes based on channel state information fed back from the decoder. In this paper, the adaptive coding scheme is employed in a mobile wireless communication system consisting of a stationary transmitter with one antenna, a wireless Rayleigh fading channel, and a mobile terminal with one or more receive antennas. The bit-error-rate at the output of the decoder is determined for various terminal speeds, time delays in the feedback channel, and number of receive antennas. The obtained results indicate that the proposed adaptive coding scheme is well suited for communications over mobile wireless channels with carrier frequencies in the high MHz range, delay spread up to 250 ns, and terminal mobility up to pedestrian speed.

I. Introduction

Many authors have studied adaptive (coded) modulation for wireless communications (see [1] and the references therein). In an earlier paper [2], we considered a general adaptive coding scheme for single-user channels with frequency-flat slowly varying multipath fading. A particular instance of this coding scheme utilizes a set of multidimensional trellis codes designed for additive white Gaussian noise (AWGN) channels of different qualities. A feedback channel makes it possible for the encoder to switch adaptively between these codes based on channel state information (CSI) fed back from the decoder, thus resulting in an overall scheme with high spectral efficiency.

The output bit-error-rate (BER) of an adaptive coding scheme may increase with growing time delay in the feedback channel and/or increasing terminal speed [3]. Since any implemented feedback channel has nonzero feedback delay, and since it is necessary to allow for mobile terminals, it is important to determine the BER degradation of the proposed adaptive coding scheme in [2]. Alouini and Goldsmith [4] have determined the BER degradation for *uncoded* adaptive modulation. In this paper, we extend their technique to determine the BER degradation of any instance of the proposed adaptive coding scheme.

We first introduce, in Section II, a mobile wireless channel with Rayleigh fading, where the mobile terminal has multiple receive antennas whose signals are combined using the *maximal ratio combining* (MRC) method [5, Ch. 5]. For any instance of the adaptive coding scheme and any number of receive antennas, Section III then shows how to determine the BER degradation associated with a nonzero feedback delay and a

nonzero terminal speed. As an example, Section IV evaluates a specific adaptive encoder and decoder (codec) utilizing a set of four-dimensional trellis codes. A conclusion is drawn in Section V.

II. System Model and Coding Scheme

The system model consists of a stationary transmitter/receiver (transceiver), a wireless frequency-flat fading channel, and a mobile transceiver, or terminal. It is assumed that the distance between the stationary transceiver and the mobile terminal is not more than a few hundred meters. We will only consider the flow of user information on the downlink. Hence, in our model the *feedback channel* (or uplink) from the terminal to the receiver will only be used for CSI.

The stationary transceiver has one transmit antenna, while the mobile terminal has H (≥ 1) receive antennas. Each of the H antenna branches is modeled as a Rayleigh fading channel with ideal coherent detection. It is assumed that the branch signals are statistically independent.

Denoting the transmitted complex baseband signal at time index $t \in \{0, 1, 2, \dots\}$ by $x(t)$, the received signal at antenna $h \in \{1, 2, \dots, H\}$ can then be written as $y_h(t) = \alpha_h(t) \cdot x(t) + n_h(t)$. Here, the stationary and ergodic *fading envelope* $\alpha_h(t)$ is a real-valued random variable with a Rayleigh distribution, and $n_h(t)$ is complex-valued AWGN with statistically independent real and imaginary components. The total one-sided power spectral density of the AWGN is denoted N_0 [W/Hz] and the one-sided channel bandwidth is denoted B [Hz].



Geir E. Øien (36) received his MSc and PhD degrees from the Norwegian University of Science and Technology (NTNU) in 1989 and 1993, respectively. From 1994 until 1996 he was with Stavanger University College as associate professor. Since 1996 he has been with NTNU, since 2001 as full professor of information theory.

Prof. Øien is a member of IEEE and the Norwegian Signal Processing Society. He is author/co-author of more than 40 research papers. His current research interests are in information theory, communication theory, and signal processing, with emphasis on analysis and design of wireless communication systems.

oien@tele.ntnu.no

Let S [W] denote the constant average transmit power. The instantaneous received *carrier-to-noise* ratio (CNR) on antenna branch h at time index t is then

$$\gamma_h(t) = \frac{\alpha_h^2(t) \cdot S}{N_0 B}, \quad h = 1, 2, \dots, H,$$

with expectation $E[\gamma_h(t)] = \bar{\gamma}_h = \Omega S / (N_0 B)$

where $\Omega = E[\alpha_h^2(t)]$ is assumed independent of h . Thus, $\bar{\gamma}_h$ is also equal for all h .

The mobile terminal implements an MRC combiner to process the H received branch signals [5, p. 316]. Since the branch signals are statistically independent, the instantaneous CNR at the output of the H -branch MRC combiner is given by $\gamma = \sum_{h=1}^H \gamma_h$.¹⁾ If we denote $E[\gamma] = \bar{\gamma}$ then $\bar{\gamma}_h = \bar{\gamma} / H$, and the *gamma* probability density function (pdf) of the instantaneous CNR γ at the output of the MRC combiner may be written as [5, Eq. (5.2-14)]

$$p_\gamma(\gamma) = \left(\frac{H}{\bar{\gamma}}\right)^H \frac{\gamma^{H-1}}{(H-1)!} \exp\left(-H\frac{\gamma}{\bar{\gamma}}\right), \quad (1)$$

$$\gamma \geq 0.$$

It is convenient to view the combination of the H antenna branches and the MRC combiner as a single channel. The instantaneous CNR γ at the output of this channel determines the *channel state* at a given time. We assume that the mobile terminal has perfect knowledge of γ . The range $[0, \infty)$ of possible CNR values is divided into $N + 1$ non overlapping intervals (or *fading regions*). At any given time the CNR will fall in one of these fading regions, and the associated CSI, i.e. the region index $n \in \{0, 1, \dots, N\}$, is sent to the stationary receiver via the feedback channel, which is assumed to be error free.

Assume that $\gamma \in [0, \gamma_1)$ in fading region 0, $\gamma \in [\gamma_n, \gamma_{n+1})$ in region $n \in \{1, 2, \dots, N-1\}$, and $\gamma \in [\gamma_N, \infty)$ in region N . Also, assume that the BER must never exceed a target maximum BER_0 . When $\gamma \in [\gamma_n, \gamma_{n+1})$ we use a multidimensional trellis code, denoted code $n \in \{1, 2, \dots, N\}$, designed to achieve a $\text{BER} \leq \text{BER}_0$ on an AWGN channel of CNR $\gamma \geq \gamma_n$. For $\gamma < \gamma_1$, i.e. γ in fading region 0, the channel conditions are so bad that no information is transmitted, and we have an *outage* during which the information flow is buffered.

Let $4 \leq M_1 < M_2 < \dots < M_N$ denote the number of symbols in N quadrature amplitude modulation (QAM) constellations of growing size, and let

code n be based on the constellation with M_n symbols. For some small fixed $L \in \{1, 2, \dots\}$, the encoder for code n accepts $L \cdot \log_2(M_n) - 1$ information bits at each time index $k = L \cdot t \in \{0, L, 2L, \dots\}$ and generates $L \cdot \log_2(M_n)$ coded bits. The coded bits specify L modulation symbols in the n th QAM constellation. These symbols are transmitted at time indices $k, k+1, \dots, k+L-1$. The L two-dimensional symbols can be viewed as one $2L$ -dimensional symbol, and for this reason the code is said to be a $2L$ -dimensional trellis code. In practice, the N codes are chosen such that they may be encoded and decoded by the same codec [2].

To determine the values of the fading region boundaries (or thresholds) γ_n , we need to determine the BER performance of each code. When code n is operating on an AWGN channel of CNR γ , the BER-CNR relationship for varying γ may be approximated by the expression

$$\text{BER} \approx a_n \cdot \exp\left(\frac{-b_n \gamma}{M_n}\right), \quad (2)$$

where $a_n (> 0)$ and $b_n (> 0)$ are constants which depend only on the weight distribution of the code [2]. These constants can be found for any given code by least-squares curve fitting of data from AWGN channel simulations to (2). The fitting must be done separately for each code in the set.

Plots of BER found in the literature indicate that the approximation in (2) is accurate for any CNR γ resulting in $\text{BER} \approx 10^{-1}$ (see Figure 1 for an example). Unfortunately, for the minimum value $\gamma = 0$, the approximation reduces to $\text{BER} \approx a_n$, and since a_n can be larger than one, (2) may be of little use for low CNRs. When we only want to approximate the BER at moderate-to-high CNRs, as was done in [2], this is not a problem. However, we need to approximate the BER for any CNR $\gamma \geq 0$ in this paper, and we will therefore use the following BER expression for code n

$$\text{BER}_n = \begin{cases} a_n \cdot \exp\left(-\frac{b_n \gamma}{M_n}\right), & \gamma \geq \gamma_n^* \\ \frac{1}{2}, & \gamma < \gamma_n^* \end{cases} \quad (3)$$

Here, the boundary

$$\gamma_n^* = \frac{\ln(2a_n)M_n}{b_n}$$

¹⁾ We suppress the time dependence from now on for notational simplicity.

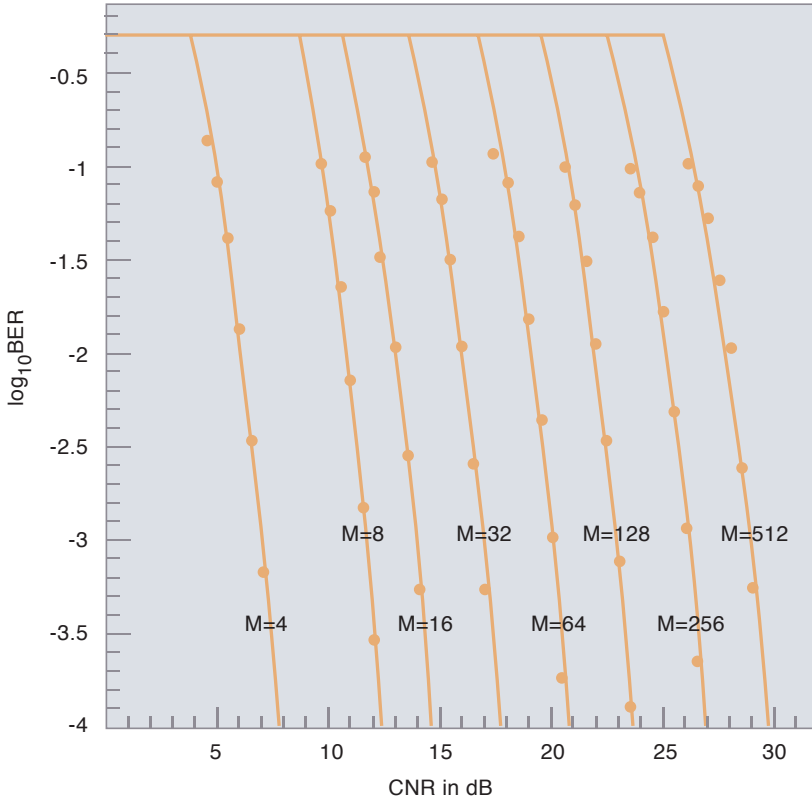


Figure 1 The boxes are BER estimates generated by software simulation and the curves are estimates obtained from (3). The labels indicate the number of symbols in the QAM signal constellations utilized by the 4-dimensional trellis codes

is the smallest CNR such that the BER is no larger than 0.5. The boundary was obtained by assuming equality in (2), setting $\text{BER} = 0.5$, and solving for γ .

For a true BER between 10^{-1} and 0.5 the exponential expression in (3) tends to produce a larger value than the true BER, assuming that the coded communication system manages to maintain synchronization. In practice, it is difficult to maintain synchronization for a very high BER, and the approximation $\text{BER} = 0.5$ may therefore be close to the true BER of a real system. If the coded system should exhibit a $\text{BER} > 0.5$ for a very low CNR, then all decoded information bits may be flipped to achieve a $\text{BER} < 0.5$. Hence, 0.5 is a reasonable upper bound on the BER.

Assuming a target BER_0 such that $\gamma > \gamma_n^*$ and setting BER_n equal to BER_0 in (3), the thresholds are given by [2]

$$\gamma_n = (M_n K_n) / b_n, \quad n = 1, 2, \dots, N \quad (4)$$

$$\gamma_{N+1} = \infty$$

where $K_n = -\ln(\text{BER}_0 / a_n)$.

The probability that γ falls in fading region n , $P_n = P(\gamma_n \leq \gamma < \gamma_{n+1})$, is given by [4, Eq. (10)]

$$P_n = \frac{\Gamma\left(H, \frac{H\gamma_n}{\bar{\gamma}}\right) - \Gamma\left(H, \frac{H\gamma_{n+1}}{\bar{\gamma}}\right)}{(H-1)!} \quad (5)$$

where

$$\Gamma(v, \mu) = \int_{\mu}^{\infty} t^{v-1} e^{-t} dt \quad (6)$$

is the complementary incomplete gamma function [6, Eq. (8.350.2)]. Since H is an integer in (5), the function may be calculated using [6, Eq. (8.352.2)].

III. BER Degradation

The BER degradation due to nonzero feedback delay and nonzero terminal speed is determined in this section. It is assumed that the communication system utilizes a set of N trellis codes with known parameters a_n and b_n .

Let the total feedback delay, τ [s], be the time between the moment the mobile terminal acquires a set of L modulation symbols and the moment the stationary transmitter activates a new code. The total feedback delay is determined by the sum of three delays: i) the processing time needed by the terminal to estimate the instantaneous CNR γ and to determine in which fading region n the CNR falls, ii) the time needed to feed back the region index n to the transmitter, and iii) the processing time needed by the transmitter to activate code n .

In a real system, the processing delay i) depends on the technique used to estimate the instantaneous received CNR, whereas the processing delay iii) depends on the encoder complexity. Since the distance between the stationary transmitter and the mobile terminal is assumed to be no more than a few hundred meters, the transmission delay ii) is mainly determined by the communication protocols.

The size of the signal constellation $M_n = M_n(\gamma)$ at time index t is a function of the instantaneous received CNR γ , but the constellation is used at time $t + \tau$ when γ has changed to γ_{τ} . Consequently, while the CNR γ falls in some fading region n , i.e. $\gamma_n \leq \gamma < \gamma_{n+1}$, the CNR γ_{τ} may fall outside this region. Substituting γ_{τ} for γ in (3), we can write the BER as a function of γ_{τ} for a given γ :

$$\text{BER}_n^{\tau}(\gamma_{\tau} | \gamma) = \begin{cases} a_n \cdot \exp\left(-\frac{b_n \gamma_{\tau}}{M_n(\gamma)}\right), & \gamma_{\tau} \geq \gamma_n^* \\ \frac{1}{2}, & \gamma_{\tau} < \gamma_n^* \end{cases} \quad (7)$$

The average BER for γ in fading region n is now given by

$$\langle \text{BER} \rangle_n^{\tau} = \int_{\gamma_n}^{\gamma_{n+1}} \int_0^{\infty} \text{BER}_n^{\tau}(\gamma_{\tau} | \gamma) p_{\gamma_{\tau} | \gamma}(\gamma_{\tau} | \gamma) d\gamma_{\tau} p_{\gamma}(\gamma) d\gamma, \quad (8)$$

where $p_\gamma(\gamma)$ is given by (1). Furthermore,

$p_{\gamma_\tau|\gamma}(\gamma_\tau|\gamma)$ is the pdf of γ_τ conditioned on γ [4]

$$p_{\gamma_\tau|\gamma}(\gamma_\tau|\gamma) = \frac{H}{(1-\rho)\bar{\gamma}} \left(\frac{\gamma_\tau}{\rho\gamma} \right)^{(H-1)/2} \cdot I_{H-1} \left(\frac{2H\sqrt{\rho\gamma\gamma_\tau}}{(1-\rho)\bar{\gamma}} \right) \cdot \exp \left(-\frac{H(\rho\gamma + \gamma_\tau)}{(1-\rho)\bar{\gamma}} \right). \quad (9)$$

The function $I_{H-1}(\cdot)$ in (9) is the $(H-1)$ th-order modified Bessel function of the first kind [7, Ch. 9]. The pdf also contains the channel power correlation coefficient ρ at lag τ . It is shown in Appendix A that ρ is given by the square of the zeroth-order Bessel function of the first kind [7, Ch. 9],

$$\rho = J_0^2(2\pi f_D \tau), \quad (10)$$

for any number of receive antennas. Here, $f_D = v/\lambda$ [Hz] is the maximum Doppler frequency shift defined by the terminal speed v [m/s] and the wavelength λ [m] of the carrier.

Using (7), the average BER in fading region n given by (8) can be rewritten as the difference between two double integrals,

$$\langle \text{BER} \rangle_n^\tau = \mathcal{I}(n) - \mathcal{J}(n),$$

where

$$\mathcal{I}(n) \stackrel{\text{def}}{=} \int_{\gamma_n}^{\gamma_{n+1}} \left\{ \int_0^\infty a_n \cdot \exp \left(-\frac{b_n \gamma_\tau}{M_n} \right) p_{\gamma_\tau|\gamma}(\gamma_\tau|\gamma) d\gamma_\tau \right\} \cdot p_\gamma(\gamma) d\gamma \quad (11)$$

and

$$\mathcal{J}(n) \stackrel{\text{def}}{=} \int_{\gamma_n}^{\gamma_{n+1}} \left\{ \int_0^{\gamma_n^*} \left[a_n \cdot \exp \left(-\frac{b_n \gamma_\tau}{M_n} \right) - \frac{1}{2} \right] p_{\gamma_\tau|\gamma}(\gamma_\tau|\gamma) d\gamma_\tau \right\} \cdot p_\gamma(\gamma) d\gamma. \quad (12)$$

The double integral $\mathcal{J}(n)$ is zero for $\gamma_n^* = 0$, i.e. when parameters a_n and b_n result in good BER approximations for any $\gamma_\tau \geq 0$. Hence, $\mathcal{J}(n)$ may be viewed as a ‘‘correction term’’ needed when a_n and b_n are only useful for $\gamma_\tau \geq \gamma_n^* > 0$.

It is shown in Appendix B that

$$\mathcal{I}(n) = \frac{a_n}{(H-1)!} \left(\frac{H}{\bar{\gamma}} \right)^H \frac{\Gamma(H, \beta_n \gamma_n) - \Gamma(H, \beta_n \gamma_{n+1})}{(\omega_n)^H} \quad (13)$$

where

$$\beta_n = \frac{H}{\bar{\gamma}} + \frac{H\rho b_n}{HM_n + \bar{\gamma}(1-\rho)b_n} \quad (14)$$

and

$$\omega_n = \frac{H}{\bar{\gamma}} + \frac{b_n}{M_n}. \quad (15)$$

From Appendix C, we have

$$\mathcal{I}(n) = S(a_n, b_n) - S\left(\frac{1}{2}, 0\right) \quad (16)$$

for

$$S(a_n, b_n) \stackrel{\text{def}}{=} a_n \frac{(1-\rho)^H}{(H-1)!} \cdot \left\{ \sum_{j=0}^{\infty} \frac{\rho^j}{(j+H-1)! j!} \left[\frac{HM_n}{b_n(1-\rho)\bar{\gamma} + HM_n} \right]^{j+H} \cdot \gamma_{\text{inc}} \left(H+j, \left[\frac{b_n}{M_n} + \frac{H}{(1-\rho)\bar{\gamma}} \right] \gamma_n^* \right) \cdot \left[\Gamma \left(H+j, \frac{H\gamma_n}{(1-\rho)\bar{\gamma}} \right) - \Gamma \left(H+j, \frac{H\gamma_{n+1}}{(1-\rho)\bar{\gamma}} \right) \right] \right\} \quad (17)$$

where

$$\gamma_{\text{inc}}(v, \mu) = \int_0^\mu t^{v-1} e^{-t} dt \quad (18)$$

is the *incomplete gamma function* [6, Eq. (8.350.1)]. Since $H+j$ is an integer in (17), the function may be calculated using [6, Eq. (8.352.1)].

The average BER over all N codes, denoted by

$\langle \text{BER} \rangle^\tau$, is equal to the expected number of information bits in error per modulation symbol divided by the expected number of transmitted information bits per modulation symbol,

$$\langle \text{BER} \rangle^\tau = \frac{\sum_{n=1}^N i_n \langle \text{BER} \rangle_n^\tau}{\sum_{n=1}^N i_n P_n} = \frac{\sum_{n=1}^N i_n [\mathcal{I}(n) - \mathcal{J}(n)]}{\sum_{n=1}^N i_n P_n}. \quad (19)$$

Here, $i_n = \log_2(M_n) - 1/L$ is the number of information bits per modulation symbol and P_n is defined by (5). In practice, the double integral $\mathcal{J}(n)$ can only be approximated since the sum in (17) must be terminated after a definite number of terms. Since each term in the sum is positive, the termination causes the expression in (19) to become an upper bound on the BER. The tightness of the bound improves as the number of terms is increased. We will use the ten first terms in the sum of (17) in the next section.

n	M_n	a_n	b_n	γ_n [dB] \approx
1	4	188.7471	9.8182	7.7
2	8	288.8051	6.8792	12.4
3	16	161.6898	7.8862	14.6
4	32	142.6920	7.8264	17.6
5	64	126.2118	7.4931	20.8
6	128	121.5189	7.7013	23.7
7	256	79.8360	7.1450	26.9
8	512	34.6128	6.9190	29.7

Table 1 Parameters a_n and b_n for example codec and calculated thresholds γ_n [dB] for target $BER_0 = 10^{-4}$

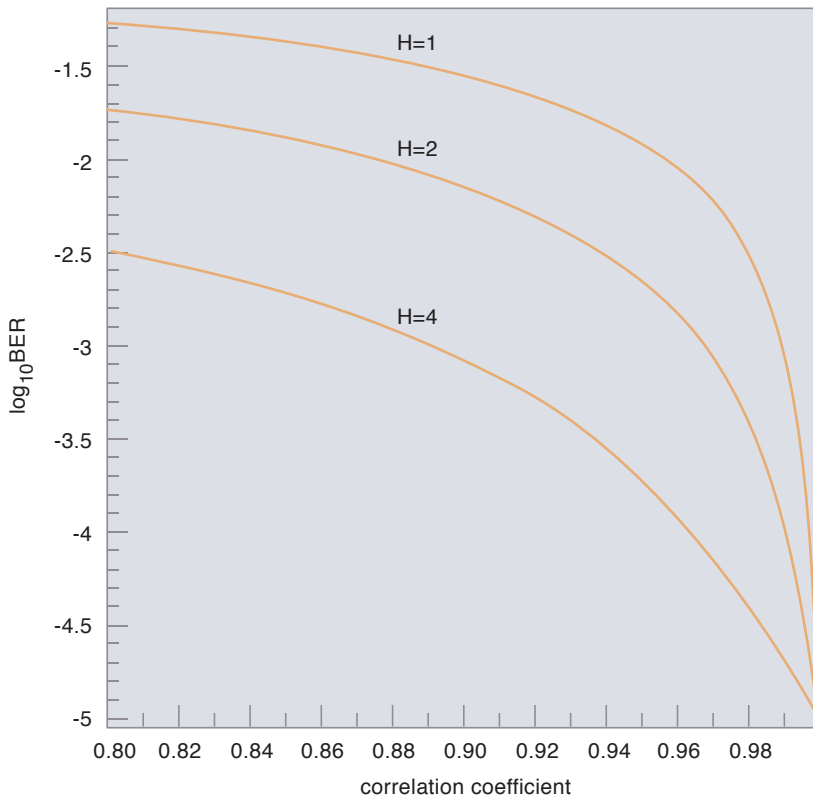


Figure 2 Base-10 logarithm of average BER for correlation coefficient $0.8 \leq \rho < 1$, target $BER_0 = 10^{-4}$, and average antenna branch CNR $\bar{\gamma}_h = 20$ dB

H	min. ρ	τ_{\max} [ms] \approx	τ_{\max}/T [symp.] \approx
1	0.997	2.7	1,080
2	0.991	3.4	1,360
4	0.963	6.9	2,760

Table 2 Minimum correlation coefficient ρ needed to achieve $BER \leq 10^{-4}$ for average antenna branch CNR $\bar{\gamma}_h = 20$ dB and different number H of receive antennas. Maximum tolerable delay τ_{\max} [ms] and number of modulation symbols transmitted during τ_{\max} for carrier frequency 1900 MHz, bandwidth 400 kHz, and terminal speed $v = 1$ m/s

IV. Evaluation of Example Codec

An adaptive codec with eight 4-dimensional trellis codes was described in [2]. The individual codes' BER performances on an AWGN channel were simulated for various CNRs. The obtained BER points (represented by boxes) are shown in Figure 1. Curve fitting with the least squares method was used to obtain the parameters a_n and b_n listed in Table 1. The corresponding BER approximations (3) are plotted in Figure 1. The expression in (4) was used to determine the tabulated thresholds²⁾ γ_n (rounded to one decimal digit) for target $BER_0 = 10^{-4}$.

Using the thresholds γ_n , setting $L = 2$, and $\bar{\gamma}_h = 20$ dB, the base-10 logarithm of the average BER (19) is plotted as a function of the correlation coefficient ρ in Figure 2 for $H \in \{1, 2, 4\}$ receive antennas. We observe that because the thresholds are chosen according to (4), the instantaneous BER is smaller than the target BER_0 for $\gamma_n < \gamma < \gamma_{n+1}$ and ρ close to one. As a result, the average BER will be below BER_0 for large ρ (see Figure 2).

Let τ_{\max} denote the maximum total delay, or *maximum tolerable delay*, for a given target BER_0 . The expression (10) for ρ can be used to determine the maximum tolerable delay τ_{\max} for different Doppler shifts f_D and targets BER_0 . The minimum values of ρ (rounded to three decimal digits) needed to achieve $\langle BER \rangle^T \leq BER_0 = 10^{-4}$ are listed in Table 2 for $H \in \{1, 2, 4\}$.

If we let the carrier frequency be $f = 1900$ MHz and use the value $c = 3 \cdot 10^8$ m/s for the speed of light, then the wavelength of the carrier frequency is $\lambda = c/f = 3/19$ ($\oplus 0.16$) m. A mobile terminal with (pedestrian) speed $v = 1$ m/s then has Doppler shift $f_D = v/\lambda = 19/3$ ($\oplus 6.33$) Hz. The corresponding maximum tolerable delays τ_{\max} (rounded to one decimal digit) are listed in Table 2.

To see that the fading is nearly constant over many hundred modulation symbols for communications at pedestrian speed, we calculate the number of symbols transmitted during the maximum tolerable delay τ_{\max} . We first need to determine a bandwidth B for which it is reasonable to assume that the fading is frequency-flat. The (rms) *delay spread*, σ_d [s], measures how much a signal component may be delayed during transmission [8, Sec. 2.2.2]. The reciprocal of the delay spread provides a measure of the width of the band of frequencies which are similarly affected by the channel response. The channel

²⁾ The thresholds in Table 1 are larger than the thresholds in [2, Table 1] because we have reduced the target BER_0 from 10^{-3} to 10^{-4} . Furthermore, the path memory length of the Viterbi decoder was set to 9 in [2] while a path memory length of 16 was used in this paper.

is therefore approximately frequency-flat if the bandwidth $B \ll 1/\sigma_d$.

At 1900 MHz, the multipath delay spread is up to $\sigma_d = 250$ ns for a cordless phone in indoor and outdoor environments [9]. Hence, we may assume that a channel with bandwidth at least up to $B = 400$ kHz has frequency flat fading. The time needed to transmit one symbol at the Nyquist signaling rate is $T = 1/B = 2.5$ μ s, resulting in τ_{\max}/T symbols being transmitted during the maximum tolerable delay. Using the rounded values of τ_{\max} in Table 2, we obtain the τ_{\max}/T values listed in the rightmost column of Table 2 for terminal speed $v = 1$ m/s and $H \in \{1, 2, 4\}$.

V. Conclusion

It has been shown (see Figure 2) that the BER performance may degrade considerably as ρ decreases, which – for a given carrier frequency – corresponds to increasing the terminal speed. However, the degradation can be mitigated by the use of MRC antenna diversity. Still, our results indicate that adaptive coded modulation may be best suited for systems with moderate mobility requirements, with terminals moving at pedestrian speed.

Appendix A – Calculation of ρ

In this appendix we show that the channel power correlation coefficient ρ is given by the expression in (10). The instantaneous received CNR on the channel may be expressed as $\gamma = \alpha^2 \cdot K$ where α^2 is the channel power gain and $K = S/(N_0 B)$. Since $E[\gamma] = \sum_{h=1}^H E[\gamma_h] = HK\Omega$, we have $E[\alpha^2] = H\Omega$. Assume that α^2 is the power gain at some time t and let α_τ^2 be the power gain at time $t + \tau$ for $\tau > 0$. The correlation coefficient ρ between α^2 and α_τ^2 is then given by

$$\rho = \frac{\text{cov}(\alpha^2, \alpha_\tau^2)}{\sqrt{\sigma_{\alpha^2}^2 \sigma_{\alpha_\tau^2}^2}} = \frac{E[\alpha^2 \alpha_\tau^2] - E[\alpha^2]E[\alpha_\tau^2]}{\sigma_{\alpha^2} \sigma_{\alpha_\tau^2}}. \quad (20)$$

The channel gain α^2 is gamma distributed [8, p. 48]. Hence, assuming that the channel power gains α^2 and α_τ^2 have the same expectations and standard deviations, we have $E[\alpha^2]E[\alpha_\tau^2] = (H\Omega)^2$ and $\sigma_{\alpha^2} \sigma_{\alpha_\tau^2} = (E[\alpha^2])^2/H = H\Omega^2$ in (20).

To calculate $E[\alpha^2 \alpha_\tau^2]$, we first compare two different expressions for the instantaneous received CNR γ . When the communication channel is viewed as a Rayleigh fading channel with a H -branch MRC combiner, then

$\gamma = \sum_{h=1}^H \gamma_h = K \sum_{h=1}^H \alpha_h^2$ where α_h^2 is the power gain on the H th antenna branch. Since we also have $\gamma = \alpha^2 \cdot K$, it follows that $\alpha^2 = \sum_{h=1}^H \alpha_h^2$ and we can write

$$E[\alpha^2 \alpha_\tau^2] = E \left[\left(\sum_{h=1}^H \alpha_h^2 \right) \left(\sum_{i=1}^H \alpha_{i,\tau}^2 \right) \right] = \sum_{h=1}^H E[\alpha_h^2 \alpha_{h,\tau}^2] + \sum_{h=1}^H \sum_{i \neq h}^H E[\alpha_h^2 \alpha_{i,\tau}^2]. \quad (21)$$

Furthermore, because the signals on different antenna branches ($h \neq i$) are statistically independent, the covariance

$$\text{cov}(\alpha_h^2, \alpha_{i,\tau}^2) = E[\alpha_h^2 \alpha_{i,\tau}^2] - E[\alpha_h^2] E[\alpha_{i,\tau}^2] = 0,$$

or equivalently, $E[\alpha_h^2 \alpha_{i,\tau}^2] - E[\alpha_h^2] E[\alpha_{i,\tau}^2] = 0$. The expression in (21) is then equal to

$$E[\alpha^2 \alpha_\tau^2] = HE[\alpha_h^2 \alpha_{h,\tau}^2] + H(H-1)\Omega^2,$$

and the correlation coefficient in (20) reduces to

$$\rho = \frac{E[\alpha_h^2 \alpha_{h,\tau}^2] - \Omega^2}{\Omega^2} \quad (22)$$

Observe that (22) is independent of the number of receive antennas H . In fact, (22) defines the correlation coefficient for a Rayleigh fading channel (without MRC). It is shown in [8, Eq. (2.68)] that the numerator in (22) is equal to $\Omega^2 J_0^2(2\pi f_D \tau)$, and as a result, ρ is given by the expression in (10).

Appendix B – Evaluation of $\mathcal{I}_1(n)$

In the following we calculate the double integral in (11). For the inner integral, BER_n in (3) is fixed since the CNR γ is fixed. It follows from (3) that

$$M_n = -\frac{b_n \gamma}{\ln(\text{BER}_n/a_n)}.$$

Using this expression for M_n and setting $D_n = -\ln(\text{BER}_n/a_n)$, the inner integral in (11) is equal to

$$\begin{aligned} \mathcal{I}_1(n, \gamma) &\stackrel{\text{def}}{=} a_n \int_0^\infty \frac{H}{(1-\rho)\bar{\gamma}} \left(\frac{\gamma_\tau}{\gamma\rho} \right)^{(H-1)/2} \\ &\cdot \exp \left(-\frac{H(\rho\gamma + \gamma_\tau)}{(1-\rho)\bar{\gamma}} - \frac{D_n \gamma_\tau}{\gamma} \right) \\ &\cdot I_{H-1} \left(\frac{2H\sqrt{\rho\gamma\gamma_\tau}}{\bar{\gamma}(1-\rho)} \right) d\gamma_\tau. \end{aligned} \quad (23)$$

Introducing the constant

$$x = \frac{\rho H^2 \gamma^2}{\bar{\gamma}(1-\rho)(H\gamma + \bar{\gamma}(1-\rho)D_n)}$$

and making the substitution

$$z = \left(\frac{H}{\bar{\gamma}(1-\rho)} + \frac{D_n}{\gamma} \right) \gamma \tau,$$

the integral (23) can be written as

$$\begin{aligned} \mathfrak{I}(n, \gamma) &= a_n \left(\frac{H\gamma}{H\gamma + \bar{\gamma}(1-\rho)D_n} \right)^H \\ &\cdot \exp \left(-\frac{\rho D_n H \gamma}{H\gamma + \bar{\gamma}(1-\rho)D_n} \right) \\ &\cdot \int_0^\infty \left(\frac{z}{x} \right)^{(H-1)/2} e^{-z-x} I_{H-1} \left(2\sqrt{xz} \right) dz. \end{aligned} \quad (24)$$

The value of the integral in (24) is equal to $Q_H(x, 0)$ where $Q_H(\cdot, \cdot)$ is the *generalized Marcum Q-function of order H* [7, Eq. (11.63)]. Since $Q_H(x, 0) = Q_1(x, 0) = 1$ for all x , we have

$$\begin{aligned} \mathfrak{I}(n, \gamma) &= a_n \left(\frac{H\gamma}{H\gamma + \bar{\gamma}(1-\rho)D_n} \right)^H \\ &\cdot \exp \left(-\frac{\rho D_n H \gamma}{H\gamma + \bar{\gamma}(1-\rho)D_n} \right). \end{aligned} \quad (25)$$

The double integral in (11) can now be written as

$$\mathfrak{I}(n) = \mathcal{F}(\gamma_n) - \mathcal{F}(\gamma_{n+1}) \quad (26)$$

for

$$\mathcal{F}(\xi) = \int_\xi^\infty \mathfrak{I}(n, \gamma) p_\gamma(\gamma) d\gamma.$$

To calculate $\mathcal{F}(\xi)$, we first observe that BER_n in (3) is no longer a constant since γ varies. Using the connection $D_n = -\ln(\text{BER}_n / a_n) = (b_n \gamma) / M_n$, it follows from (1) and (25) that

$$\begin{aligned} \mathcal{F}(\xi) &= \frac{a_n}{(H-1)!} \left(\frac{H}{\bar{\gamma}} \right)^H \left(\frac{HM_n}{HM_n + \bar{\gamma}(1-\rho)b_n} \right)^H \\ &\cdot \int_\xi^\infty \gamma^{H-1} \exp(-\beta_n \gamma) d\gamma \end{aligned}$$

where β_n is defined by (14). Substituting $t = \beta_n \gamma$ and observing that

$$\left(\frac{HM_n}{HM_n + \bar{\gamma}(1-\rho)b_n} \right)^H \left(\frac{1}{\beta_n} \right)^H = \left(\frac{1}{\omega_n} \right)^H$$

for ω_n defined by (15), we get

$$\mathcal{F}(\xi) = \frac{a_n}{(H-1)!} \left(\frac{H}{\bar{\gamma}} \right)^H \frac{\Gamma(H, \beta_n \xi)}{(\omega_n)^H} \quad (27)$$

where $\Gamma(\cdot, \cdot)$ is given by (6). The expression for $\mathfrak{I}(n)$ in (13) is now obtained from (26) and (27).

Appendix C – Evaluation of $\mathfrak{I}(n)$

We shall calculate the double integral $\mathfrak{I}(n)$ defined by (12). We first split the double integral in two to obtain

$$\begin{aligned} \mathfrak{I}(n) &= \int_{\gamma_n}^{\gamma_{n+1}} \left\{ \int_0^{\gamma_n^*} a_n \cdot \exp \left(-\frac{b_n \gamma \tau}{M_n} \right) p_{\gamma_\tau | \gamma}(\gamma_\tau | \gamma) d\gamma_\tau \right\} \\ &\cdot p_\gamma(\gamma) d\gamma \quad (28) \\ &- \int_{\gamma_n}^{\gamma_{n+1}} \left\{ \int_0^{\gamma_n^*} 2^{-1} p_{\gamma_\tau | \gamma}(\gamma_\tau | \gamma) d\gamma_\tau \right\} p_\gamma(\gamma) d\gamma. \end{aligned} \quad (29)$$

The second integral (29) is a special case of the first integral (28) with $a_n = 2^{-1}$ and $b_n = 0$.

Hence, we only need to consider the first integral. The pdf $p_{\gamma_\tau | \gamma}(\gamma_\tau | \gamma)$ defined by (9) contains the $(H-1)$ th-order modified Bessel function of the first kind defined by [7, Eq. (9.28)]

$$I_\nu(z) = \left(\frac{1}{2} z \right)^\nu \sum_{j=0}^{\infty} \frac{\left(\frac{1}{2} z \right)^{2j}}{(j+\nu)! j!}$$

for ν an integer. Using this definition, the inner integral in (28) is equal to

$$\begin{aligned} &a_n \left[\frac{HM_n}{b_n(1-\rho)\bar{\gamma} + HM_n} \right]^H \cdot \exp \left(-\frac{H\rho\gamma}{(1-\rho)\bar{\gamma}} \right) \\ &\sum_{j=0}^{\infty} \frac{1}{(j+H-1)! j!} \left[\frac{M_n H^2 \rho \gamma}{(1-\rho)\bar{\gamma} \{ b_n(1-\rho)\bar{\gamma} + HM_n \}} \right]^j \\ &\gamma_{\text{inc}} \left(H + j, \left[\frac{b_n}{M_n} + \frac{H}{(1-\rho)\bar{\gamma}} \right] \gamma_n^* \right) \end{aligned}$$

where $\gamma_{\text{inc}}(\cdot, \cdot)$ is defined by (18). The outer integral in (28) is then equal to (17). Since the double integral in (29) is a special case of the double integral in (28), the “correction term” $\mathfrak{I}(n)$ is now given by (16).

References

- 1 Hole, K J, Øien, G E. Spectral efficiency of adaptive coded modulation in urban micro-cellular networks. *IEEE Trans. Veh. Technol.*, 50 (1), 205–222, 2001.
- 2 Hole, K J, Holm, H, Øien, G E. Adaptive multidimensional coded modulation over flat fading channels. *IEEE J. Select. Areas Commun.*, 18 (7), 1153–1158, 2000.
- 3 Goeckel, D L. Adaptive coding for time-varying channels using outdated fading estimates. *IEEE Trans. Commun.*, 47 (6), 844–855, 1999.

- 4 Alouini, M-S, Goldsmith, A J. Adaptive M-QAM modulation over Nakagami fading channels. *Proc. 6th Communications Theory Mini-Conference (CTMC VI)* in conjunction with IEEE Global Communications Conference (GLOBECOM'97), Phoenix, Arizona, Nov. 1997, 218–223.
- 5 Jakes, W C (ed.). *Microwave Mobile Communications*. NJ, Piscataway, IEEE Press, second ed., 1994.
- 6 Gradshteyn, I S, Ryzhik, I M. *Table of Integrals, Series, and Products*. San Diego, CA, Academic Press, fifth ed., 1994.
- 7 Temme, N M. *Special Functions – An Introduction to the Classical Functions of Mathematical Physics*. New York, NY, John Wiley, 1996.
- 8 Stüber, G L. *Principles of Mobile Communication*. Norwell, MA, Kluwer, 1996.
- 9 Ue, T et al. Symbol rate and modulation level-controlled adaptive modulation/TDMA/TDD system for high-bit-rate wireless data transmission. *IEEE Trans. Veh. Technol.*, 47 (4), 1134–1147, 1998.

Shannon Mappings for Robust Communication

TOR A. RAMSTAD



Tor A. Ramstad (60) received his *Siv.Ing.* and *Dr.Ing.* degrees in 1968 and 1971, respectively, both from the Norwegian University of Science and Technology (NTNU). He has held various positions in the Dep. of Telecommunications at NTNU, where since 1983 he has been a full professor of telecommunications. In 1982–83 he was a visiting associate professor at the University of California, Santa Barbara; he was with the Georgia Institute of Technology in 1989–90 as a visiting adjunct professor, and again at UCSB as a visiting professor in 1997–98. Dr. Ramstad's research interests include multirate signal processing, speech and image processing with emphasis on image and video communications, where joint source-channel coding is a central topic
tor@tele.ntnu.no

Shannon's geometric interpretation of messages and channel representations in communicating time-discrete, amplitude-continuous source symbols is exploited in this paper. The basic idea is to map what we can call *source space* symbols into *channel space* symbols, where the two symbols possibly have different dimensions. If we decrease the dimension through this operation, bandwidth reduction is obtained. If, on the other hand, the dimension is increased when mapping from the source to the channel space, bandwidth expansion results. In practical systems several mappings are applied for different dimension changes depending on the importance of the source symbols as well as the available channel resources. The paper also discusses theoretical limits for Gaussian sources and channels, expressed by OPTA (optimal performance theoretically attainable), and compares practical mapping constructions with these limits. Finally, the paper demonstrates how efficient image communication systems can be designed, and shows that much higher robustness towards channel variations can be obtained than is possible for pure digital systems.

1 Shannon Mappings

Modern telecommunications are to a large degree based on the works of Nyquist and Shannon. The present work is no exception, but it is specifically related to a less known general idea presented by Shannon in his famous 1949 paper [33]. There he suggests mapping time discrete signals from a continuous multidimensional source space to a continuous channel space of different dimensionality. By this mechanism one can obtain *bandwidth compression*, which will increase the number of simultaneous users on a physical channel, or *bandwidth expansion* if it is necessary to increase the received signal fidelity due to a noisy channel.

When we initially started the work presented in this article, we were not aware of Shannon's idea. When we discovered that Shannon had introduced the geometrical mappings in his 1949 paper [33], we were pleased and assumed we were on the right track. To honor Shannon, we would like to propose to call the mappings used in this paper *Shannon mappings*. Shannon did not pursue his idea, partly because technology was not ready for designing and implementing such systems at the time. Today the situation is quite different.

The mapping description is very general and leads to a deeper understanding of the general communication problem. One can cast such techniques as quantization and modulation into this framework. The term joint source-channel coding is often encountered in modern literature. The mapping idea can probably incorporate all the different techniques suggested for source-channel coding, and it is probably the most general way of looking at the entire communication problem.

In this article we will present the general ideas, introduce some simple examples of how to construct mappings for special sources and channels, and show complete systems for image transmission.

2 Introduction to the Communication Problem

The basic objective in efficient communication of natural signals like speech, images, and video, can be stated as: optimizing for conveying as many signals as possible with a specified quality over a given physical channel. The channel is usually band-limited, it has certain power and/or amplitude constraints, e.g. due to battery lifetime and regulations for radio transmission, and the channel will usually distort the signal by insertion of different noise types, like thermic noise or contamination by crosstalk from other electrical signals, or linear and nonlinear distortions due to non-ideal behavior of the channel.

The exploitation of a given physical channel can be influenced by different measures. These include signal adaptation to the channel characteristics, driving power, relay amplification/regeneration, and optimal receivers. Given the channel capacity, the number of useful signals that can be transmitted is also influenced by signal compression. However, the amount of compression is limited by the unavoidable noise generated when amplitude-continuous signals are compressed.

Under certain simplifying, but rather realistic conditions, the *channel capacity* can be calculated [33]. To approach the channel capacity, which guarantees lossless transmission of multi-level signals, one has to resort to systems of infinite complexity, which also require infinite

delay. In such a channel, the compression operation can be viewed as a separate problem [32, 34] due to Shannon's *separation theorem*. To obtain a minimum digital representation, infinite complexity is also required for the codec.

The aim of this paper, however, is to contribute to developing understanding for the possible gains which can be reached by doing joint compression and channel coding, when we constrain the allowed complexity and delay. We stress that very good robustness is obtained for the proposed systems without resorting to explicit error protection.

2.1 Discrete Transmission over Noise-free, Band-limited Channels

Throughout the discussion we assume that all useful signals are either ideally band-limited, or can be made close to that through lowpass filtering without loss of subjective quality. (The pass-band of the filter can be freely chosen.) This enables us to sample the signal at or above the Nyquist rate without loss of information. If the bandwidth is B , and the minimum necessary sampling frequency is $F_s = 2B$, then the sampled signal can be transmitted over an ideal and noise-free *Nyquist channel* with bandwidth $W = B$. That is, the time discrete signal requires the same bandwidth as the original analog signal.

In practice, we need to relax this assumption somewhat. First of all, some oversampling is necessary to allow for realistic low-pass anti-aliasing filters, and also, any Nyquist channel needs some roll-off factor in order to obtain finite order filters. We can account for both effects by including some oversampling factor α implying that the practical channel bandwidth must be $W = \alpha B$. α is a design parameter that can be made close to 1 by increasing the filter complexities.

2.2 Limits in Signal Communication

The above discussion on error free transmission specifies the number of samples that can be transmitted for a given channel bandwidth. It does not include limitations induced by noise, which is an omnipresent phenomenon. As we shall see, noise-free transmission would imply infinite channel capacity!

The second important aspect in communications is how signals carry information and how they can be represented by samples that can be conveyed through the physical channel. Actually, from a mathematical point of view any analog signal contains infinite information, it is only when accepting noise contamination that finite information results. However, this is not a disaster, as we shall see that the human observer is quite tolerant towards certain degradations.

2.2.1 Distortion and Rate

Natural signals conveying meaningful information to humans invariably have inherent sample-to-sample dependencies. Such dependencies are usually characterized as *redundancy*. Furthermore, human perception is not capable of distinguishing signals with certain distortions from the pure original. Even observable distortion is tolerated. Most communication systems rely heavily on the user tolerance to distortions. The amount of imperceptible distortion is often called *irrelevancy*. The amount of acceptable distortion can be called *noise tolerance*.

To assess distortion we need distortion measures. Such measures should agree as much as possible with the perception of the receiver. When for natural signals the receiver is a human being, the distortion measure should mimic human perception. This turns out to be rather complex both for visual and auditory perception. Although it does not reflect human perception well, the most common distortion measure is the *mean squared error (mse)*. If the original- and reconstructed signals are given by vectors with components $x_i, i = 1, 2, 3, \dots, M$ and $\hat{x}_i, i = 1, 2, 3, \dots, M$, respectively, the *mse* is defined by

$$D = \varepsilon \left\{ \frac{1}{M} \sum_{i=1}^M (x_i - \hat{x}_i)^2 \right\}, \quad (1)$$

where $\varepsilon\{ \}$ is the expectation operator.

If we have a good signal model, we can, according to Shannon, evaluate the necessary rate for obtaining a given distortion. A Gaussian source of independent and identically distributed samples has a simple rate distortion (R-D) function measured in bits/sample

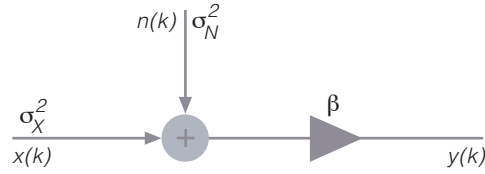
$$R = \begin{cases} \frac{1}{2} \log_2 \left\{ \frac{\sigma_X^2}{\sigma_D^2} \right\}, & \text{for } \sigma_D^2 \leq \sigma_X^2, \\ 0, & \text{for } \sigma_D^2 > \sigma_X^2, \end{cases} \quad (2)$$

where σ_X^2 is the signal power and σ_D^2 is the accepted distortion. As the rate can never be negative, the limiting case means that the noise is set equal to the signal. That is, it is preferable to transmit nothing for this case. The expression σ_X^2 / σ_D^2 is called the *signal-to-noise ratio (SNR)*, and is usually measured in dB.

With correlated sources, the R-D function is more complex. The main point is that due to the interrelation between the samples, a correlated signal can be coded at a lower rate for the same distortion.

Although we measure the rate in bits/sample, the formula does not imply that we must quantize

Figure 1 Optimal system with no bandwidth alteration



the signal. This is only a measure of information, that is, it measures the information in a signal when a certain noise level is acceptable.

2.2.2 Channel Capacity

The Nyquist theorem tells us how many samples can be transmitted given the bandwidth. It does not tell us how much information can be conveyed by each sample when the channel has a given signal-to-noise ratio. Shannon's channel capacity provides this measure.

The simplest case is when the channel is memoryless and power limited, that is, the channel power is given by σ_C^2 , and the channel is corrupted by white, Gaussian noise with power σ_N^2 . Then the channel has a capacity measured in bits per sample:

$$C = \frac{1}{2} \log_2 \left\{ 1 + \frac{\sigma_C^2}{\sigma_N^2} \right\}. \quad (3)$$

σ_C^2 / σ_N^2 is called the *channel signal-to-noise ratio* (CSNR), and is also usually measured in dB.

It is again more complicated to find the capacity of channels with other constraints and memory as well as other noise types. For our purpose this channel will illustrate our main points.

2.2.3 Optimal Performance Theoretically Attainable (OPTA)

OPTA is a measure for the limits for efficient communications. It can be derived from the rate-distortion function and the channel capacity.

The simplest case occurs when we transmit one source sample per channel sample. This means that the source rate has to be equal to the channel capacity

$$\begin{aligned} R_s &= \frac{1}{2} \log_2 \left\{ \frac{\sigma_X^2}{\sigma_D^2} \right\} = C \\ &= \frac{1}{2} \log_2 \left\{ 1 + \frac{\sigma_C^2}{\sigma_N^2} \right\} \end{aligned} \quad (4)$$

which implies that

$$\frac{\sigma_X^2}{\sigma_D^2} = \left(1 + \frac{\sigma_C^2}{\sigma_N^2} \right). \quad (5)$$

This means that the signal-to-noise ratio in the received signal is given by the signal-to-noise ratio on the channel plus 1.

An interesting fact is that unlike almost all other cases, the OPTA can be reached using a very simple implementation.

A possible system model is given by Figure 1. The signal samples are transmitted without any modification over a Nyquist channel with additive Gaussian noise. We reach the OPTA condition for this system when the signal is also Gaussian by selecting

$$\beta = \frac{\sigma_X^2}{\sigma_X^2 + \sigma_N^2}. \quad (6)$$

Remember that $\sigma_C^2 = \sigma_X^2$ for the above case.

If we try to construct an optimal system based on the *separation theorem* for this case, we would have to apply an infinite-dimensional vector quantizer for the source coder followed by a channel coder with infinite delay.

In the above case the signal bandwidth and the channel bandwidth are the same. If, for some reason, this is not acceptable, there must be a sample rate change. This is easily understood if we consider rate per time unit rather than rate per symbol. The source produces $2B$ samples per second, which indicates that its rate is given by $2BR_s$ bits per second. Likewise, the channel can transmit $2W$ symbols per second, which gives a capacity of $2WC$ bits per second. Setting these two rates equal; we obtain OPTA for the general case (still assuming the same simple source and channel models).

$$2B \frac{1}{2} \log_2 \left\{ \frac{\sigma_X^2}{\sigma_D^2} \right\} = 2W \frac{1}{2} \log_2 \left\{ 1 + \frac{\sigma_C^2}{\sigma_N^2} \right\}. \quad (7)$$

Solving this equation with respect to the source signal-to-noise ratio, we obtain

$$\frac{\sigma_X^2}{\sigma_D^2} = \left(1 + \frac{\sigma_C^2}{\sigma_N^2} \right)^{W/B} \quad (8)$$

We observe that the bandwidth relation W/B enters the equation. This bandwidth change can be obtained only by dimension change where M source samples are combined into K channel samples, where $K/M \approx W/B$.

In Figure 2 the OPTA-curves for different sample compression ratios are given. We have included dimension increase in the plots, which gives sample rate expansion, as well as dimension reduction. We mentioned initially that the

ultimate aim is to make rate or bandwidth reduction, but this can be obtained only if the channel has a sufficient CSNR for the required received signal quality. If the channel is not good enough even without compression, sample rate increase is necessary.

Practical signal sources are usually decomposed into sub-sources with different variances and even time variations. Each source would normally require different SNR, implying that some sub-sources can be compressed, others can be sent unmodified, while some have to be expanded. Actually, some sub-sources are insignificant and can be skipped altogether. This is a natural consequence of the condition that $R_s = 0$ for $\sigma_D^2 = \sigma_X^2$ in Equation 2. The average rate required by all these sub-sources is the interesting number for assessing the efficiency of the system.

This theory gives the limits for any single-source single-channel system, but it does not tell us how to get close to these limits. The rest of this paper will provide examples of possible methods where nonlinear mappings are used for dimension change. In most of these methods no bit representation is used, but it is convenient to compare the results to more traditional results where bits give an intermediate representation form.

3 Geometric Description of Dimension Change

Define the input signal as a vector \mathbf{x} consisting of M components. The vector can be represented geometrically as a point in a Euclidean space of dimension M . Mathematically we say that $\mathbf{x} \in \mathbb{R}^M$. Then the vector components x_1, x_2, \dots, x_M are the coordinate values of the vector. A dimension change can be performed by a *mapping* Ψ from the space of dimension M to a space of dimension K ($\Psi : \mathbb{R}^M \rightarrow \mathbb{R}^K$) as

$$\mathbf{y} = \Psi(\mathbf{x}). \quad (9)$$

The operator Ψ is generally nonlinear, but we can occasionally make use of linear operators.

Mappings without dimension change ($K = M$), and mappings with dimension increase ($K > M$) can be made invertible, whereas dimension reducing mappings ($K < M$) for all practical purposes involve approximative operations, and are thus not invertible.

It is instructive to split the mapping operation into two parts; an *approximation operation* q , which maps the complete space onto a subspace, and the dimension-changing operation T , which is an invertible operation:

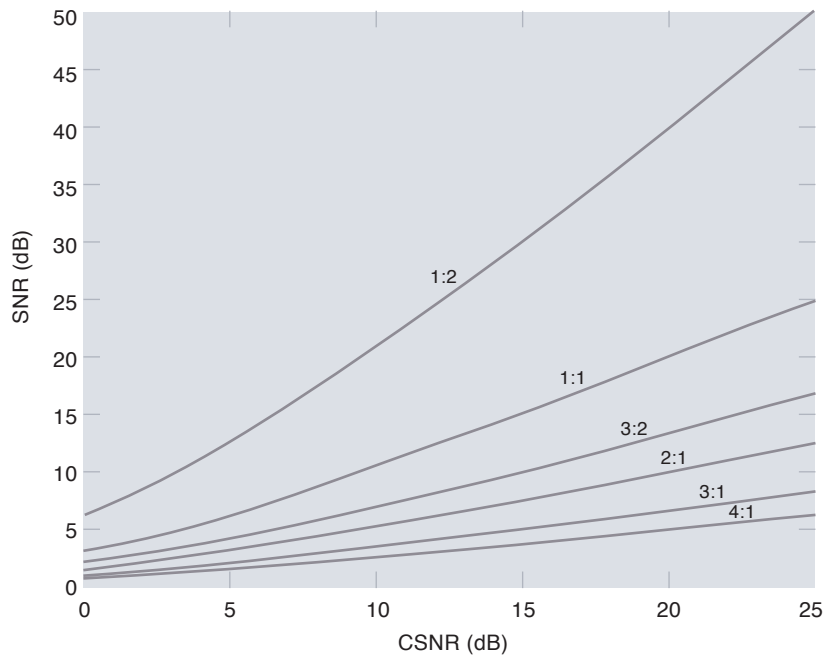


Figure 2 OPTA curves with M/K as parameter

$$\Psi = q \circ T. \quad (10)$$

The operations are illustrated in Figure 3. The operation q is, of course, an identity operation when Ψ is invertible.

The simplest way of doing dimension reduction is to skip some of the vector components.

Although dimension increase can always be done without approximation, the most common way of increasing the signal dimension is by quantization, which is not an invertible operation.

3.1 Quantization in Source-to-Channel Mappings

Quantization must be applied for making a digital representation from an analog signal. We will discuss the most general form called *vector quantization*, and show that the principle is very simple and closely related to the mappings discussed in this paper.

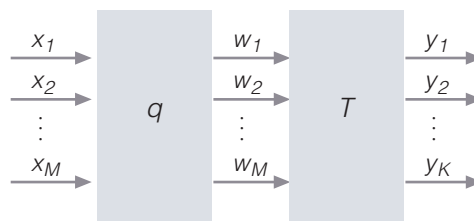


Figure 3 Mapping from higher to lower dimension as a two-stage process. q indicates approximation while T represents the mapping between dimensions

A vector quantizer can be split into an approximation part q and a part which maps to a discrete representation. This is an intermediate representation from which we can map to the channel in many ways.

q is a mapping of $\mathbf{x} \in \mathbb{R}^M$ to a finite set $\mathbf{W} = \{\mathbf{w}_0, \dots, \mathbf{w}_{L-1} | \mathbf{w}_i \in \mathbb{R}^M\}$ of representation vectors:

$$q : \mathbb{R}^M \rightarrow \mathbf{W}. \quad (11)$$

The mapping q can be defined on a *partition* of \mathbb{R}^M into L non-overlapping, M -dimensional *cells* $\{C_i\}$ according to the rule

$$\mathbf{x} \in C_i \Rightarrow q(\mathbf{x}) = \mathbf{w}_i, \quad i = 0, 1, \dots, L-1, \quad (12)$$

where $\{C_i\}$ satisfies

$$\bigcup_{i=0}^{L-1} C_i = \mathbb{R}^M \text{ and } C_i \cap C_j = \Delta \text{ for } i \neq j. \quad (13)$$

In a VQ setting the collection of representation vectors is referred to as the *codebook*. The cells C_i , called *Voronoi regions*, can be thought of as solid polygons in the M -dimensional space \mathbb{R}^M .

The index i identifies that vector uniquely and is therefore an efficient representation of the vector. The vector can be reconstructed approximately as $\mathbf{x} \approx \mathbf{w}_i$ by looking up the representation vector in cell number i . Thus, the bit rate in bits per sample in this scheme is $b = \log_2(L)/M$. A further bit reduction can also be obtained by entropy coding of the indices, provided the symbols have different probabilities.

The next step is to design the channel representation. The simplest example is transmitting the indices bitwise on a binary channel using $K = b$. We then obtain a mapping $\Psi : \mathbb{R}^M \rightarrow \mathbb{R}^{\log_2(L)}$. Whether this is a bandwidth expansion or reduction depends on whether $\log_2(L) \gtrless M$.

With reference to Figure 3 the dimension-changing mapping T consists of the *index assignment* and the channel representation of the index.

But there are many other ways of making channel representations of the index. One could group the bits in the index into, say b_1 bits, and transmit the message by 2^{b_1} -level signals. This would reduce the bandwidth by a factor of b_1 . Often we want to protect the bits using forward error correction (FEC), in which case we add parity bits, thus increasing the necessary channel bandwidth.

Let us analyze the simplest system of all where we use scalar, uniform quantization ($M = 1$) with b bits per sample and binary signaling on the

channel. Uniform quantization requires that all Voronoi regions are of length Δ on the real line.

The reconstructed signal can then be written

$$\hat{x} = \Delta \text{sign}(x) \sum_{i=1}^{b-1} a_i 2^{-i}, \quad (14)$$

where the a_i 's represent the bits received with values 0 and 1, and b is the total number of bits when we reserve one bit for the sign of x .

Consider now the result of a bit error when the a_i 's are transmitted directly on the binary channel. An error in the least significant bit hardly influences the reconstruction, whereas a bit error in the most significant bit changes the sign of the signal and thus can change the value greatly if the amplitude is large at the same time. This problem relates to what Shannon calls the *threshold effect* (see below). The advantage of this method is that we require only a modest CSNR to get a very low bit error rate.

If we, on the other hand, transmit the complete index directly using $L = 2^b$ channel levels, a transition from one state to another only makes a change in the least significant bit. This requires a much higher CSNR than for binary transmission, but at the same time guarantees a moderate channel noise influence on every sample.

Even though we can represent the process of going from the real line via quantization to some channel representation in many steps, the two important steps are the approximation step which maps the real line to a discrete subset on the real line, and the mapping from these discrete values to a discrete channel space of possibly higher dimension.

We claimed earlier that dimension increase does not have to involve any approximation. And as a matter of fact, the quantization step is not necessary. We will return to this shortly.

We will now leave methods based on Shannon's separation theorem and study mappings in greater detail and present some important examples. We start with a discussion of methods suggested by Shannon.

3.2 Ideas in Shannon's Paper

Shannon, in his 1949 paper [33], suggests a mapping which is reproduced in Figure 4. If used for signal expansion, the length along the curve from some reference point represents the source sample amplitude, while the two coordinates of any point on the curve gives the two channel samples, which can be represented as a QAM symbol. Channel noise will disturb the

signal. If, in the receiver, the decision device projects the noise corrupted sample down to the closest point on the curve, then small noise samples will only move the signal along the curve and thus produce insignificant change. If the channel noise reaches a certain level, there is a probability of crossing to a neighboring line. This is what Shannon refers to as the *threshold effect*. This effect will be encountered for all non-linear expanding mappings.

Shannon also mentions that the same mapping can be used for signal compression. Then the two coordinates represent the signal vector, and are approximated by the closest point on the curve. The channel symbol can be selected as the distance along the curve from a reference point to the approximation point on the curve.

We will demonstrate that Shannon-like mappings work quite well for signal expansion, and argue that they may also perform well for compression if the signal is uniformly distributed over the square.

Shannon suggested another algorithm for signal compression. For the special case of 2:1 mapping ($\Psi : \mathbb{R}^2_+ \rightarrow \mathbb{R}^1_+$) he first represents the two components in decimal or binary form:

$$\begin{aligned} x_1 &= 0.c_1c_2c_3 \dots \\ x_2 &= 0.d_1d_2d_3 \dots, \end{aligned} \quad (15)$$

where c_i and d_i , $i = 1, 2, 3 \dots$ are the digits. The one-dimensional signal y is obtained by picking digits from x_1 and x_2 alternately and inserting them into y as

$$y = 0.c_1d_1c_2d_2c_3d_3 \dots, \quad (16)$$

In this way it is possible to obtain an exact representation of any two real numbers by one real number using an infinite number of digits. In a communication system y can be transmitted as one sample.

It is important to note that if x_1 and x_2 are both described by b bits, then y needs $2b$ bits to be exactly represented. Assume we transmit first the original samples using one multilevel symbol for each component. This would require 2^b levels for each symbol. If y is transmitted as a multilevel symbol, the required bandwidth is reduced by a factor of 2, but the number of levels now needs to be 2^{2b} for exact representation. For transmission over noisy channels this would require approximately twice the signal-to-noise ratio measured in dB for transmitting y compared to transmitting the original symbols x_1 and x_2 . So if we define a bandwidth-SNR product (Hz \times dB), it remains unchanged for this mapping operation.

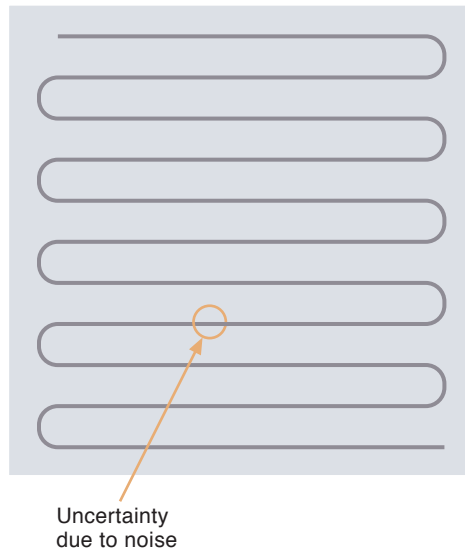


Figure 4 One-to-two-dimensional mapping suggested by Shannon

It is easy to generalize this method to any mapping $\Psi : \mathbb{R}^M \rightarrow \mathbb{R}^1$ or even $\Psi : \mathbb{R}^M \rightarrow \mathbb{R}^K$.

In order to compare this system with others, we cast the algorithm into geometric form. Assume that the two components have a support $x_k \in [0, 1]$, $k = 1, 2$, which is indicated by the way we have written the numbers in Equations 15 and 16. Also represent the components with

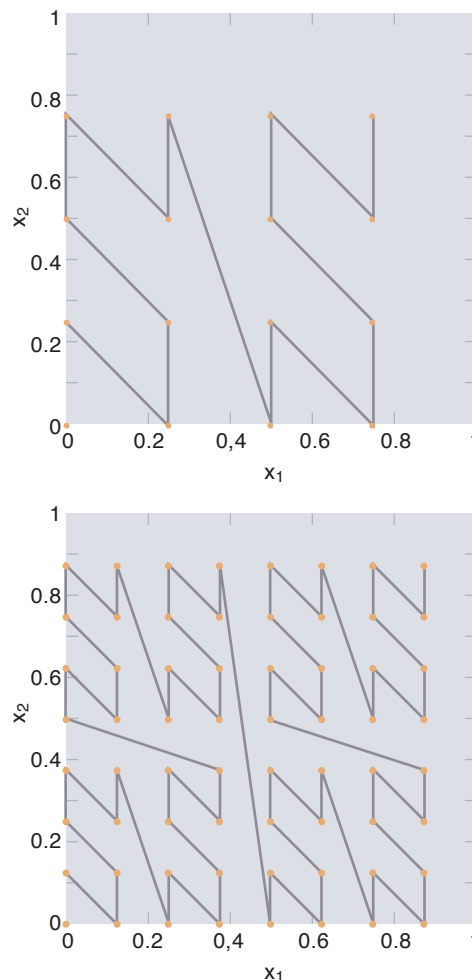


Figure 5 Cantor maps for signal dimension change representing the mapping given by Equations 15 and 16

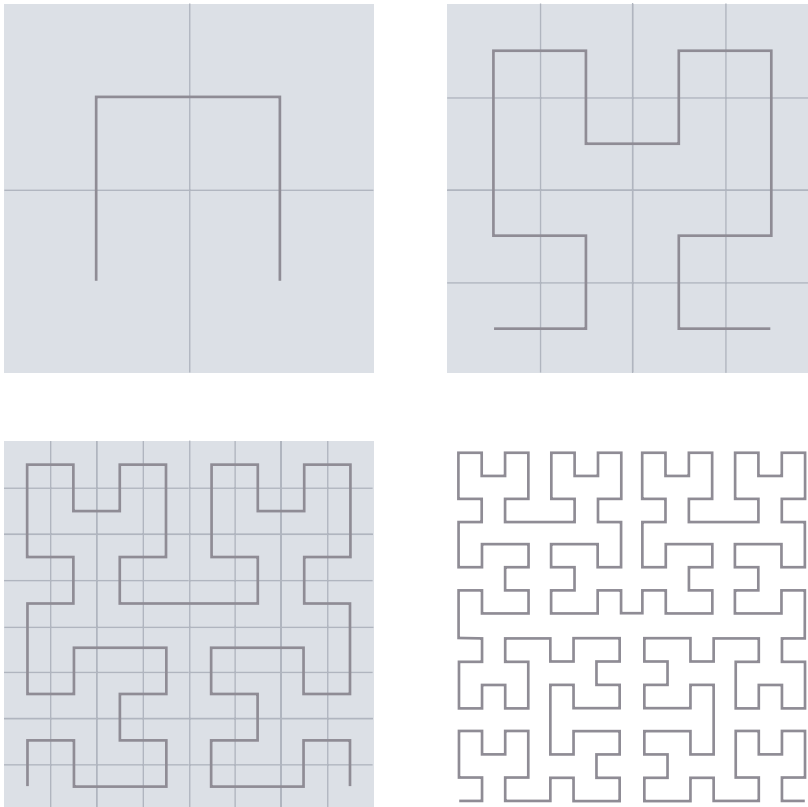
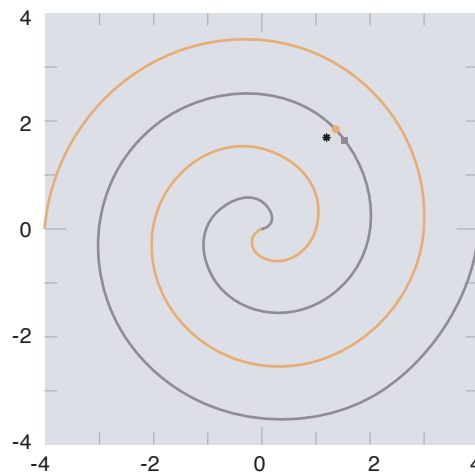


Figure 6 The first 4 iterations for constructing a space-filling Hilbert-curve

a finite number of bits, which obviously is the same as quantization. Then mappings using 3 and 4 bits are shown in Figure 5 in the upper and lower parts, respectively.

The lines connecting the quantization values indicate how y is constructed. The numeric value of y is proportional to the number of nodes which is traversed from the origin to the actual vector point. From the two figures one easily deduces a way to construct systems with more bits in a systematic way. In the limit when $b \rightarrow \infty$, the whole space will be densely populated by points, and thus the accuracy of the compressed signal is also guaranteed.

Figure 7 Double spiral of Archimedes. The orange spiral can be represented by positive channel symbols, while the grey spiral can be represented by negative channel symbols. The star represents a 2-D input vector. The channel representation is obtained by approximation to the closest point on the spiral indicated by a circle. The channel adds noise to the channel sample, which in turn causes the representation point to move along the spiral, as indicated by the diamond



In the next two sections we will dig deeper into the geometric construction of mappings for dimension change in communication systems.

3.3 Dimension Reduction

Let us discuss further how we can make exact representations of a signal vector by using a vector of a lower dimension. We consider 2:1 mappings which try to represent the 2-D space, which is a plane, by use of a one-dimensional continuous curve. This is somewhat different from the previous example where we used a discrete subset as an intermediate representation. This new problem is related to what is called space-filling curves, or Peano-curves after the inventor.

Assume that the region of support for the 2-D signal is a unit square. Figure 6 shows how the classical Hilbert-curve can be constructed to fill the entire space. To be able to represent any point exactly, an infinite number of iterations has to be used.

A possible one-dimensional representation of any point in the two-dimensional space is the distance from the entrance-point of the curve to the point on the curve which coincides with the coordinates of the vector. It is obvious that this distance will, in general, be infinite. To cope with this problem for practical use, we must resort to a limited number of iterations, which would mean that we can only represent most points of the space *approximately*, but the distance to all points will then be finite.

3.3.1 Signal Approximation

We can conclude that it is necessary to make approximations when we lower the signal dimensionality and want to transmit the resulting signal components over a channel with finite signal-to-noise ratio.

The subspace generated by q in Figure 3 must have certain topological properties suited for subsequent mapping to the lower dimension in the operator T .

A very good continuous mapping for the case when the region of support is a disk of radius 1, is a double spiral composed of two spirals of Archimedes. The spirals can be described parametrically as,

$$\begin{aligned} x_1 &= 2\Delta \frac{\theta}{2\pi} \cos(\theta), \\ x_2 &= 2\Delta \frac{\theta}{2\pi} \sin(\theta), \end{aligned} \quad (17)$$

and

$$\begin{aligned} x_1 &= 2\Delta \frac{\theta}{2\pi} \cos(\theta + \pi), \\ x_2 &= 2\Delta \frac{\theta}{2\pi} \sin(\theta + \pi). \end{aligned} \quad (18)$$

θ is the rotation angle to the point $[x_1, x_2]$ relative the curves' derivative at the origin. $2\Delta\theta/\pi$ is the radius to the point, while Δ is the distance between two neighboring crossings of the real axis of either of the two spirals.

A spiral example is shown in Figure 7. Any 2-D vector, as e.g. the value represented by the star, will be approximated to the closest point on one of the spirals, in this case the point indicated by the circle. This is very similar to vector quantization. In VQ the approximations are points in the space, while here it is a continuous curve.

The transform T can be chosen in many ways. One possibility is to make y equal to the length from the approximation point on the spiral along the spiral to the origin. One spiral can be represented by positive amplitudes while the other can be represented by negative channel samples. Or one could select y to be the rotation angle θ .

To fill the entire space with the two curves we need to let $\Delta \rightarrow 0$. Then the spiral length will be infinite.

It is also easy to conceive systems for making space-filling curves for mappings from higher dimensions to one dimension. A one-dimensional approximate representation of a three-dimensional sphere can be visualized by a ball of yarn. The thread is the curve, and a one-dimensional representation of a point within the sphere could be the length of the thread to the closest possible point on the thread.

T can be generalized to any non-linear length adjustments, as if the thread in the ball of yarn was made from rubber, and could be stretched unevenly.

3.3.2 Channel Noise

In the complete communication chain, the channel noise is the next obstacle when we want to maximize throughput with a given fidelity. A more complete signal chain is shown in Figure 8.

In the model we represent the channel noise by the vector \mathbf{n} . At the receiver an approximate inverse operation, R , tries to minimize the noise effect while preserving the original signal as well as possible. The received signal vector is thus

$$\hat{\mathbf{x}} = R \circ (\mathbf{n} + T \circ q)(\mathbf{x}). \quad (19)$$

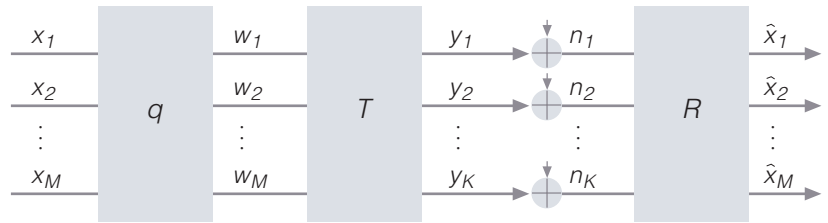


Figure 8 Overall system model

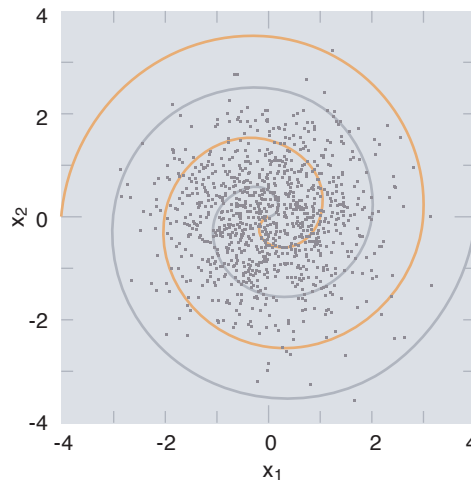


Figure 9 Scatter plots of Gaussian signals with standard deviation $\sigma_x = 1.0$ inside the spiral

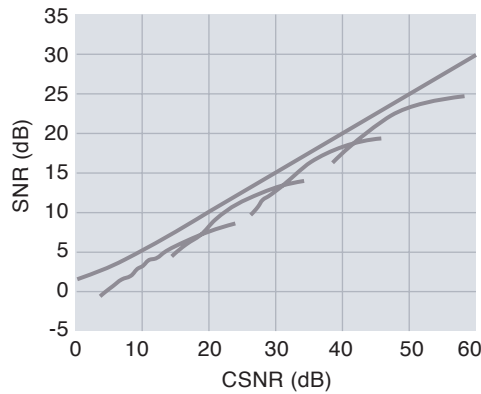
Let us try to get more insight into the noise problem by studying the spiral example further. Figure 7 shows the effect of approximation and channel noise. As the amplitude of the transmitted signal is corrupted by noise, the reconstruction will be inexact. If the channel sample was originally obtained as the length along the spiral to the approximation point, the received signal will give the length along the spiral to the decoded point as indicated in Figure 7 with a diamond.

How do the two noise contributions interact? It is obvious that the approximation noise depends on the density of the spiral relative to the signal components' standard deviation. We illustrate this in Figure 9. The noise is more severe when the standard deviation of the signal is small relative to the density of the spiral.

But the denser the spiral becomes, the larger the channel amplitudes get. If the channel power (or channel amplitudes, if channel is amplitude limited) is too large, a downscaling is necessary. In the receiver an upscaling must be applied, which also increases the noise with the upscaling factor.

We observe a typical trade-off: If we lower the approximation noise by making the spiral denser, the influence from the channel noise will become more severe, and vice versa. There

Figure 10 SNR versus CSNR for different values of $\sigma_X/\Delta = [0.1, 1.0, 2.0, 4.0]$. The OPTA-curve is also shown for the 2D-1D mapping



exists a balance between the two contributions which will make the system optimal for a given channel signal-to-noise ratio (CSNR). Figure 10 shows results from several simulations when transmitting Gaussian, white noise over a Gaussian, memoryless channel using different spirals and different channels.

It is quite interesting that the closest point from each of the simulated curves to OPTA is in the range of 1 – 2 dB. By studying the background material more closely, another interesting conclusion is that at the closest point to OPTA the relation $\sigma_N/\Delta = 0.35$ holds approximately for a large range of conditions when no scaling of the PAM symbols was performed.

Another important aspect of the double spiral approximation shown in Figure 9 is that it runs through the origin and covers the plane in a symmetric manner for the negative and positive channel amplitudes (drawn as grey and yellow lines, respectively). This implies that the channel representation is symmetric (provided that the source samples are rotation symmetrically distributed), and that the transmitted power is low because the probability density function is peaky at the origin for Gaussian signals, and will be represented by small channel amplitudes.

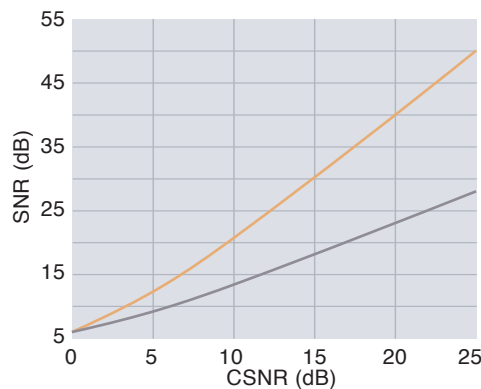


Figure 11 Comparison of the performance of a code where the samples are sent twice to the OPTA-curve for two times bandwidth expansion

We can summarize the important requirements for obtaining a good dimension reducing mapping.

1. The mapping should cover the entire space in such a way that any point is mapped to a close representation point.
2. Probable source symbols should be mapped to channel symbols with low amplitudes to minimize the average channel power.
3. Signals that are close in the channel space should remain close when mapped back to the source space. This will prevent small channel noise samples from inducing large errors in the decoded signal. The opposite is not necessary. Two close source samples may well be mapped to entirely different regions in the channel space.

Based on the above observations, will Shannon's mapping in Figure 4 perform well for dimension reduction? For the region of support shown in the figure, the curve covers the space quite nicely. If we pick the center of the figure as the reference point for zero channel amplitudes, the power requirement will also be satisfied if the probability density function is uniform over the square. On the other hand, if the signal is Gaussian, then the channel power will be higher for this mapping than for the spiral mapping. Finally, the channel noise will perturb the signal in an acceptable way. Altogether, these speculations indicate that Shannon's original mapping would perform fairly well, especially for uniformly distributed signals.

3.4 Dimension Increase

Dimension increase becomes necessary to improve the fidelity in transmitting signal amplitudes over noise prone channels.

A well-known mapping is obtained by transmitting a signal amplitude K times and averaging the output signal. This is therefore a $1 : K$ mapping. This will improve the signal-to-noise ratio in the receiver if each of the samples is contaminated by independent noise components. This is a *repetition code*.

It is easy to conclude that we gain 3 dB in SNR per doubling of K . A plot of the SNR versus CSNR using a repetition code for $K = 2$ is shown in Figure 11. There is a striking difference between the obtained performance and OPTA except at very low rates.

Let us take a closer look at the above repetition code in terms of geometry and try to find the reason for its poor performance. Figure 12 shows the repetition code when $K = 2$. Both

channel components are equal, which implies that they lie on a diagonal line in the square channel space. It is immediately clear that the channel space is not well exploited. Most of the space is empty!

3.4.1 Error-free mapping $\Psi : \mathbb{R} \rightarrow \mathbb{R}^2$

Although expanding mappings can certainly be devised for dimension change by rational ratios, we limit our discussion to the case $\Psi : \mathbb{R} \rightarrow \mathbb{R}^2$.

A very simple error-free expanding mapping uses one component that represents a discretized version of the signal and an extra component that represents the difference between the exact signal and the discretized signal.

The discretized signal can be transmitted as a PAM multilevel signal

$$y_1 = K_1 q(x), \quad (20)$$

where K_1 is a scaling factor. The second component is the correction term which can be transmitted as a continuous PAM signal

$$y_2 = K_2(x - q(x)),$$

where K_2 is a second scaling factor. The power is distributed among the two samples through the scaling factors. The quantizer is optimized both for decision and representation levels.

Two typical resulting mappings (from [2]) that have been optimized for different CSNRs are shown in Figure 13. Observe that the orange lines are not part of the mappings, but illustrate the connection between the different parts. The performance of this type of mapping is given in Figure 14. It performs much better than the repetition code presented in Figure 11!

To approach the mapping suggested by Shannon, every second of the horizontal lines in the previous mapping is reversed, resulting in the optimized system in Figure 15. The performance of this system is slightly worse than the system in Figure 13. The original mapping suggested by Shannon is thus expected to perform quite well for signal expansion.

4 Joint Source Coding and Modulation Incorporating Signal Decomposition

Real-world signals are much more complex than the signals we have studied so far. As a matter of fact, the type of information conceivable by human observers requires structures in the signal that involve statistical variations and sample dependencies. This can be modeled as signals with short-term statistics, such as spectra, that change from position to position. The non-white

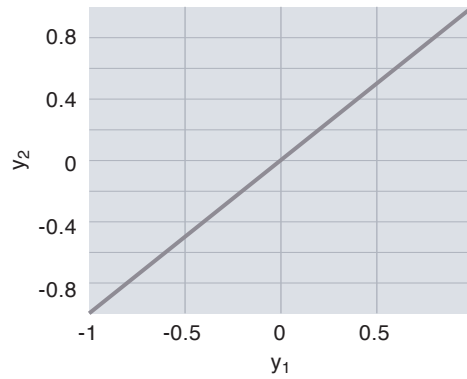


Figure 12 Repetition expansion from one to two dimensions ($\Psi : \mathbb{R} \rightarrow \mathbb{R}^2$)

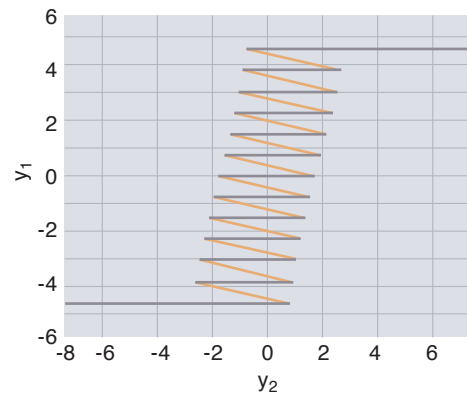


Figure 13 Optimized 1D-to-2D mapping for CSNR = 20 dB (top) and 3 dB (bottom)

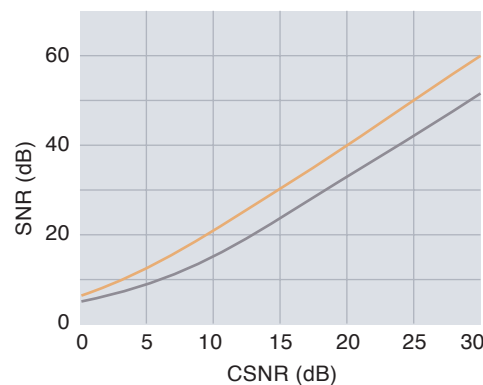
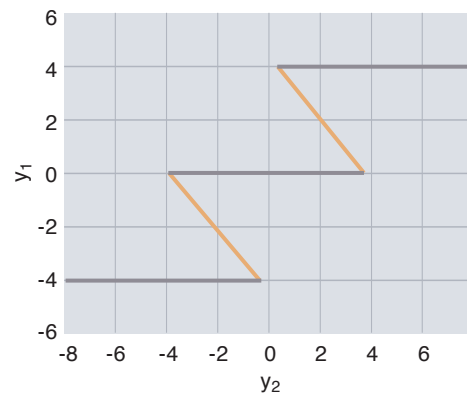
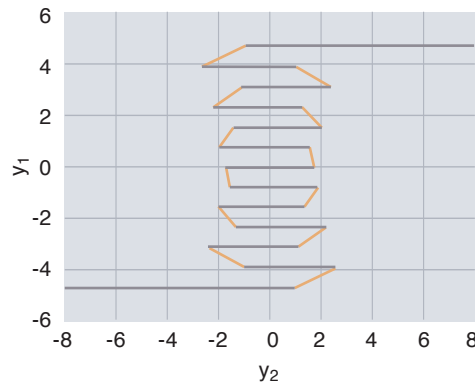


Figure 14 Performance of the 1D-to-2D mapping as a function of the CSNR when the system is optimized for each point on the curve. The upper curve is OPTA

Figure 15 “Shannon-look-alike” mapping



spectra, which account for sample dependencies, imply that some signal decomposition should be applied before any mapping takes place to decorrelate the signal. The variabilities must be accounted for by some type of adaptivity.

Signal decomposition can be performed in frequency selective filter banks. If non-overlapping bands are implemented, the output subband channels will be uncorrelated. With many channels the bands will be narrow, which implies that each band also will be close to white. The filter bank outputs have different power in each channel, which also vary with position.

Figure 16 Signal communication system including signal decomposition and mapping allocation

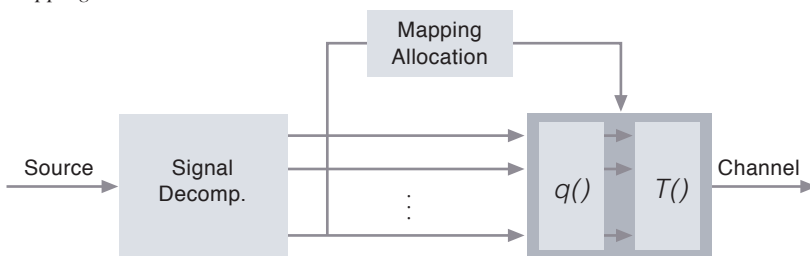
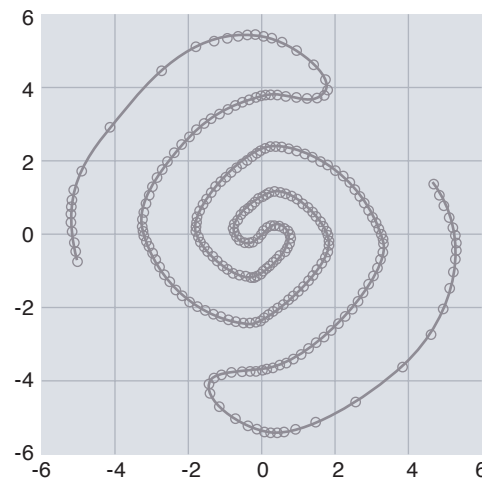


Figure 17 2:1 mapping for Laplacian source. The mapping was optimized for a white, Gaussian channel with CSNR = 23.1 dB. 256 point were used for the optimization



A complete system for signal transmission including signal decomposition is shown in Figure 16. In this system the outputs from the filter bank are analyzed and classified. The classification information is used to select different mappings that are pre-optimized for the encoder.

4.1 System Example

In the following we present an example taken from [2]. We give a brief review of the most important aspects of the system. Details can be found in the thesis.

In this system the signal decomposition is performed in a filter bank using “System K” from [1]. This is a separable filter bank which in each dimension first splits the signal into 8 subbands of equal size. The resulting low-pass band is further split dyadically into three stages. The filter coefficients were found by optimizing for coding gain.

Altogether 5 different mappings are used in the coder. These are, in terms of their $K : M$ ratio given by 1:4, 1:2, 2:3, 1:1, 2:1. The dimension reducing mappings are optimized for Laplacian sources, because this corresponds most closely to the actual distribution. The optimization method was developed by Fuldseth [6]. It optimizes a discrete set of points for minimum mse taking the approximation noise and the channel noise into account. That is, the mappings are optimized for a certain CSNR. The 1:2 mapping is shown in Figure 17. Notice that the mappings have a “spiral form”. If we drew an optimized spiral for a Gaussian source, it would look even more like Archimedes’ spirals.

The 1:1 mapping is based on Equation (6) although it is optimal only for Gaussian sources. In [2] it is shown that the deviation from OPTA is slight even when the source is Laplacian. The 1:2 mappings are of the type shown in Figure 13.

A very important part of the coder is the mapping allocation [2] included in Figure 16. This is a mechanism for using the available resources optimally. The resources in this case means *bandwidth* and *channel power*. Bandwidth is directly related to the different mapping ratios, while power is a consequence of the mappings, but can be altered by including scaling factors when inserting the signal components into the channel. In this coder the allocation is based on the local variances in each of the subbands.

Some digital side information must be conveyed to the receiver in order to inform about which mappings have been used. Full error protection is provided for this information to make sure that it is not lost before the noise destroys all mean-

ingful information anyway. Channel resources are allocated also to this part of the message.

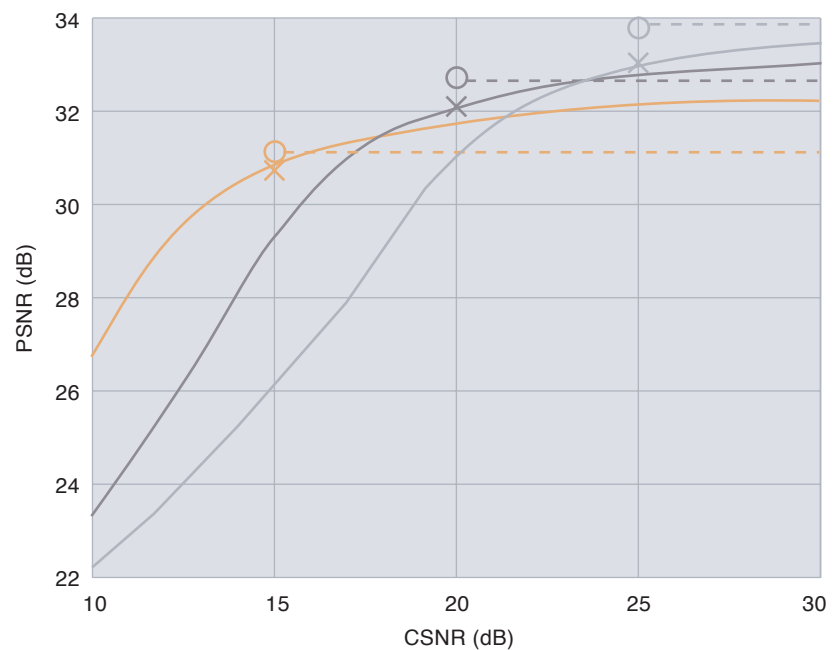
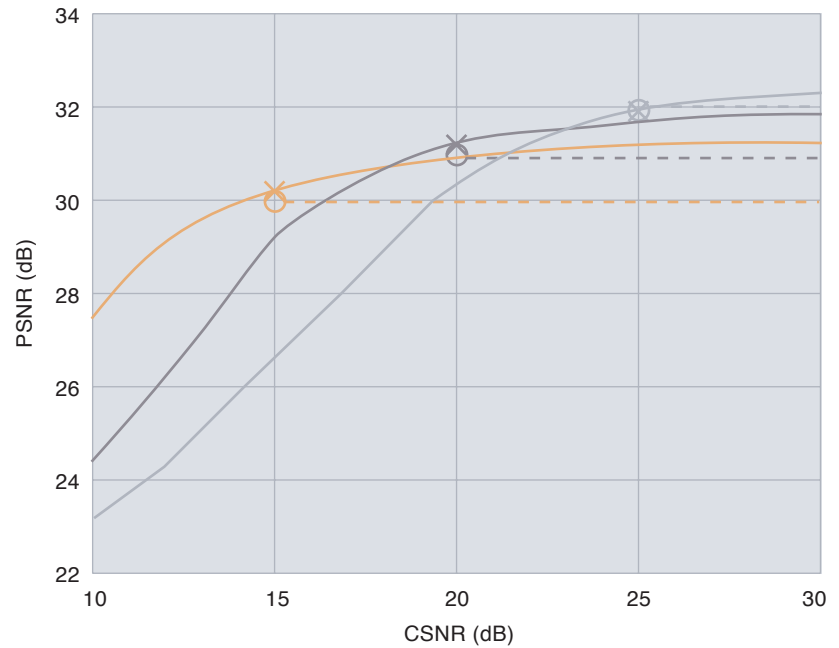
It is difficult to make a completely fair comparison to other systems. In [2] the reference system uses JPEG2000 part 1 baseline coder (ISO/IEC, 2001), implemented in verification model 8.0. Using QAM modulation with Gray coding and low-density parity check codes with soft decision, performance 3 dB away from the channel capacity can be achieved for an additive, white, Gaussian noise channel [28]. By assuming this model and error-free performance down to this CSNR, and breakdown after this point, the simulation results for two images are shown in Figure 18. PSNR means *Peak Signal-to-Noise Ratio*, and is defined as the ratio between the quadratic peak value divided by the noise power, measured in dB. It is a common quality measure for image evaluation.

Three curves are shown for each image. They represent systems optimized for the CSNR indicated by the circles and crosses. For the reference coder this is the CSNR for which the coder breaks down. This requires a certain bit rate for error protection. The system does not improve for the given design when the channel CSNR increases. This is different for the proposed system.

For the “Goldhill” image, which is quite detailed, the proposed coder outperforms the reference coder everywhere. The situation is different for “Lenna”. This image is much smoother and favors the reference coder. But the proposed coder still offers graceful degradation and can be used for lower CSNRs.

We have argued that the signal-to-noise ratio does not correspond to our perception. Therefore the visual quality of the received signal should also be studied. In the images in Figure 19 the new system is compared to the JPEG2000 system combined with the efficient channel code at CSNR = 20 dB for the two upper images. The lower image is for the new system at CSNR = 17 dB. This is at a point where the JPEG2000 system breaks down, so there is no need to show that image, as it contains only rubbish.

What clearly should be observed is that JPEG2000 blurs the tiles and the brick structure of the walls. The new system maintains more of that even at 17 dB CSNR. Notice, however, that there is a light spot on the roof of the left-hand house which is probably due to a severe channel noise component.



5 Conclusion

This paper uses the geometrical mapping method suggested by Shannon for making channel representations from signal vectors. This is a direct method which does not need the intermediate quantization step for data reduction. What we optimize for is bandwidth- and power use, which are the natural resources available. The paper shows results for an image coder recently developed. It gives comparable coding results as the JPEG2000 coder combined with the best channel coding methods available, and it offers much better robustness towards channel changes. It should also be noted that the complexity of the proposed system can be quite low, and there is no extra delay for channel coding for the main portion of the information.

Figure 18 Simulation results for “Goldhill” (upper figure) and “Lenna” (lower figure) including three systems optimized for different CSNRs (marked with star or circle). The dashed lines represent the reference coder, where the circles indicate the breakdown point. The solid lines give the performance of the proposed system. The crosses mark the design point for each curve

Figure 19 Image coding examples at the rate 0.1 channel symbols per pixel. Upper and lower images result from the proposed coder at CSNR = 20 dB and 17 dB, respectively. The middle image is by the reference coder at CSNR = 20 dB



Many other systems based on the mapping method have been devised and simulated including video coders and other channel representations, such as phase modulation.

We believe that the robustness offered, and the possible gains obtained by further development of such systems could make them good candidates for wireless systems, especially for image and video communication, but also for broadcasting.

The paper has for the most part only referenced the central Shannon papers plus the most important Dr. Ing. theses because these make the most complete description of the methods. However, there exist several papers which present parts of the methods and results, and others that present

aspects not covered in those references. For completeness the following references are therefore provided: [20, 18, 24, 25, 22, 19, 14, 23, 26, 27, 21, 8, 10, 11, 12, 13, 6, 9, 7, 16, 17, 15, 4, 5, 3, 29, 30, 31].

References

- 1 Balasingham, I. *On Optimal Perfect Reconstruction Filter Banks for Image Compression*. Trondheim, Norwegian University of Science and Technology, 1998. (PhD thesis.)
- 2 Coward, H. *Joint source-channel coding : Development of methods and utilization in image communications*. Trondheim, Norwegian University of Science and Technology, 2002. (PhD thesis.)
- 3 Coward, H, Ramstad, T A. Bandwidth doubling in combined source-channel coding of memoryless Gaussian sources. In: *Proc. IEEE Int. Symp. Intell. Signal Processing Commun. Syst. (ISPACS)*, 1, 571–576, Honolulu, HI, USA, November 2000.
- 4 Coward, H, Ramstad, T A. Quantizer optimization in hybrid digital-analog transmission of analog source signals. In: *Proc. Int. Conf. on Acoustics, Speech, and Signal Proc. (ICASSP)*, Istanbul, June 2000, 2636–2640. IEEE.
- 5 Coward, H, Ramstad, T A. Robust image communication using bandwidth reducing and expanding mappings. In: *Thirty Fourth Asilomar Conference on Signals, Systems and Computers*, 2, 1384–1388, Pacific Grove, CA, USA, October 2000.
- 6 Fuldseth, A. *Robust Subband Video Compression for Noisy Channels with Multilevel Signaling*. Trondheim, Norwegian University of Science and Technology, 1997. (PhD thesis.)
- 7 Fuldseth, A, Fischer, T R, Ramstad, T A. Channel-optimized trellis-coded vector quantization for channels with a power constraint. In: *Proc. Information Theory Workshop (ITW-98)*, San Diego, USA, February 1998.
- 8 Fuldseth, A, Lervik, J M. Combined source and channel coding for channels with a power constraint and multilevel signaling. In: *Proc. ITG Conference*, München, Germany, October 1994, 429–436. ITG.
- 9 Fuldseth, A, Ramstad, T A. Robust and efficient video communication based on combined source- and channel coding. In: *Proc. Nordic Signal Processing Symposium*

- (NORSIG -97), Tromsø, Norway, May 1995, 65–68.
- 10 Fuldseth, A, Ramstad, T A. Combined video coding and multilevel modulation. In: *Proc. Int. Conf. on Image Processing (ICIP)*, Lausanne, Switzerland, September 1996, I, 941–944.
 - 11 Fuldseth, A, Ramstad, T A. Robust subband video coding with leaky prediction. In: *Seventh IEEE Digital Signal Processing Workshop*, Loen, Norway, September 1996, 57–60.
 - 12 Fuldseth, A, Ramstad, T A. Bandwidth compression for continuous amplitude channels based on vector approximation to a continuous subset of the source signal space. In: *Proc. Int. Conf. on Acoustics, Speech, and Signal Proc. (ICASSP)*, 1997, IV, 3093–3096.
 - 13 Fuldseth, A, Ramstad, T A. Channel-optimized subband video coding for channels with a power constraint. In: *Proc. Int. Conf. on Image Processing (ICIP)*, 3, 428–431, Santa Barbara, CA, USA, October 1997.
 - 14 Grøvlen, A, Lervik, J M, Ramstad, T A. Combined digital compression and digital modulation. In: *Proc. NORSIG-95 (Signal Processing Symposium)*, Stavanger, Norway, September 1995, 69–74. IEEE/NORSIG.
 - 15 Hjørungnes, A. *Optimal bit and power constrained filter banks*. Trondheim, Norwegian University of Science and Technology, 2000. (PhD thesis.)
 - 16 Hjørungnes, A, Ramstad, T A. Linear solution of the combined source-channel coding problem using joint optimal analysis and synthesis filter banks. In: *Thirty-First Asilomar Conference on Signals, Systems and Computers*, Naval Postgraduate School, San Jose, CA, USA, 2, 990–994. Maple Press, November 1997.
 - 17 Hjørungnes, A, Ramstad, T A. On the performance of linear transmission systems over power constrained, continuous amplitude channels. In: *Proc. for the UCSB Workshop on Signal & Image Processing*, Santa Barbara, USA, December 1998, 31–35.
 - 18 Lervik, J M. Integrated system design in digital video broadcasting. *Piksel'n*, 10 (4), 12–22, 1993.
 - 19 Lervik, J M. Joint optimization of digital communication systems: Principles and practice. In: *Proc. ITG Conference*, München, Germany, October 1994, 115–122. ITG.
 - 20 Lervik, J M, Eriksen, H R, Ramstad, T A. Bandwidth efficient image transmission system based on subband coding. A possible method for HDTV. In: *Proc. NOBIM Conf.*, Lillehammer, Norway, February 1993. (In Norwegian.)
 - 21 Lervik, J M, Fischer, T R. Robust subband image coding for waveform channels with optimum power- and bandwidth allocation. In: *Proc. Int. Conf. on Acoustics, Speech, and Signal Proc. (ICASSP)*, München, Germany, April 1997. IEEE.
 - 22 Lervik, J M, Fuldseth, A, Ramstad, T A. Combined image subband coding and multilevel modulation for communication over power and bandwidth limited channels. In: *Proc. Workshop on Visual Signal Processing and Communications*, New Brunswick, NJ, USA, September 1994, 173–178. IEEE.
 - 23 Lervik, J M, Grøvlen, A, Ramstad, T A. Robust digital signal compression and modulation exploiting the advantages of analog communication. In: *Proc. IEEE GLOBE-COM*, Singapore, November 1995, 1044–1048. IEEE.
 - 24 Lervik, J M, Ramstad, T A. An analog interpretation of compression for digital communication systems. In: *Proc. Int. Conf. on Acoustics, Speech, and Signal Proc. (ICASSP)*, Adelaide, South Australia, April 1994, 5, V–281–V–284. IEEE.
 - 25 Lervik, J M, Ramstad, T A. A new approach to objective evaluation of power- and bandwidth-limited integrated communication systems. In: *Proc. Nordic Signal Processing Symposium (NORSIG -94)*, Ålesund, Norway, June 1994, 49–54. NORSIG.
 - 26 Lervik, J M, Ramstad, T A. Robust image communication using subband coding and multilevel modulation. In: *Proc. 1996 Symposium on Visual Communications and Image Processing (VCIP-96)*, Orlando, FL, USA, March 1996, SPIE 2727, 2, 524–535. SPIE/IEEE.
 - 27 Lervik, J M. *Subband Image Communication over Digital Transparent and Analog Waveform Channels*. Trondheim, Norway, Norwegian University of Science and Technology, 1996. (PhD thesis.)

- 28 Myhre, B, Markhus, V, Øien, G E. LDPC coded adaptive multilevel modulation for slowly varying Rayleigh-fading channels. In: *Proc. Norwegian Signal Processing Symp. (NORSIG)*, Trondheim, Norway, October 2001.
- 29 Ramstad, T A. Efficient and robust communication based on signal decomposition and approximative multi-dimensional mappings between source and channel spaces. In: *Proc. NORSIG*, Helsinki, Finland, September 1996.
- 30 Ramstad, T A. Digital image communication. In: *Proc. International Workshop on Circuits, Systems and Signal Processing for Communications '97*, Tampere, Finland, April 1997.
- 31 Ramstad, T A. Robust image and video communication for mobile multimedia. In: *NATO Advanced Study Institute, Signal Processing for Multimedia*, Il Cioccho, July 1998.
- 32 Shannon, C E. A mathematical theory of communication. *Bell Syst. Tech. J.*, 27, 379–423 and 623–656, 1948.
- 33 Shannon, C E. Communication in the presence of noise. *Proc. IRE*, 37, 10–21, January 1949.
- 34 Shannon, C E. Coding theorems for a discrete source with a fidelity criterion. *IRE Nat. Conv. Rec.*, March 1959, 142–163.

Information Theory

A lecture presented at a study session for radio technology and electro-acoustics
at Farris Bad, 16–18 June 1950
by Graduate Engineer Nic. Knudtzon,
Norwegian Defence Research Establishment, Bergen



Dr. Nic. Knudtzon (80) obtained his Engineering degree from the Technical University of Norway, Trondheim in 1947 and his Doctor's degree from the Technical University in Delft, the Netherlands in 1957. 1948–1949 he was with the Research Laboratory of Electronics, Massachusetts Institute of Technology, working with information theory and experiments. 1950–1955 he was with the Norwegian Defence Research Establishment, Bergen, working on the development of microwave radio links; and from 1955 to 1967 he was Head of the Communications Division at Shape Technical Center in The Hague, Netherlands, where his efforts went into the planning of military telecommunications networks and systems in Western Europe. From 1968 to 1992 he was Director of Research at the Norwegian Telecommunications Administration, working on the planning of future telecommunications systems, networks and services. Dr. Knudtzon has been member of government commissions and various committees, including the Norwegian Research Council, the National Council for Research Libraries, the International Telecommunications Union, EURESCOM, etc.

This is a translation into English of the paper "Informasjonsteori", which appeared in *Elektroteknisk Tidsskrift* 63 (30), pp. 373–380, 1950. The translation was done by Berlitz GlobalNET and final quality control was done by Geir E. Øien.

1 Introduction

Information theory is a branch of statistical communications theory. Figure 1-1 is a schematic representation of a communications system; it consists of an information source, a communications channel and a destination, the communications channel consisting of a transmitter, a connection and a receiver. The information source generates the *messages* that are to be transmitted. The transmitter converts the messages to a suitable *signal* for transmission, and in the receiver the inverse operation takes place. The destination is the person or equipment to whom, or to which, the message is addressed. On the way, noise is added. In the following treatment, we will ignore the effects of distortion of messages resulting from non-linear characteristics of the equipment, and other system "errors".

We cannot know the content of the individual messages beforehand, all we can know is the statistical characteristics of the class of messages we wish to transmit. Telecommunication is therefore a statistical process, and communications systems must be constructed for a *specified class of messages with given statistical characteristics.*^{*)} In order to make an objective assessment of the ability of a system to transmit information, it is necessary to define a unit of information, just as the volt is the unit of electrical voltage. We will now define such a unit. By means of this we will then study the information in different messages, how these should be converted to achieve an efficient transmission chan-

nel, and how much information it is possible to transmit through a channel with a given bandwidth and signal-to-noise ratio.

We are not concerned with the semantic content of the message (its meaning); a communications engineer must be capable of constructing an efficient telegraphy system for Greek, knowing the statistical characteristics of that language, even though he does not understand Greek.

We will attempt to reach our conclusions by means of clear and simple considerations, avoiding abstract mathematics (probability theory and dimension theory) which would be necessary for exact derivations. Hence, we can only consider very simple examples by means of practical applications.

2 Classification of Systems According to Message Type

We can distinguish between the following systems:

- Discrete system.* The message and the signal both consist of a series of discrete (discontinuous) symbols. An example of this is ordinary telegraphy, in which the message is a series of letters and the signal consists of dots and dashes.
- Continuous system.* The message and the signal are both continuously varying. An example is ordinary telephony, in which the message consists of pressure variations in the air

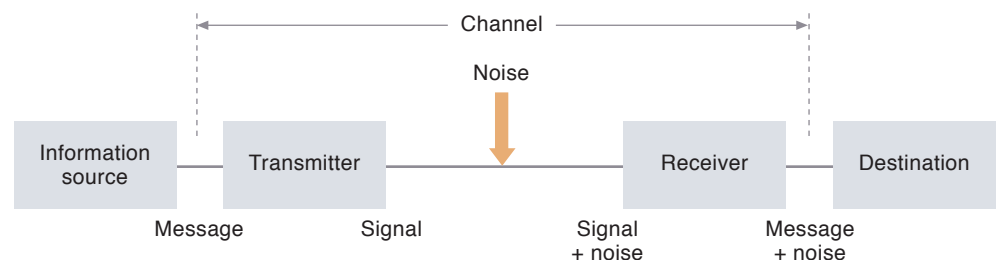


Figure 1-1 A communications system

^{*)} See also "Statistical Communications theory. A brief overview of the problem", an introductory lecture presented at a study session for radio technology and electro-acoustics, 16–18 June 1950. *Teknisk Ukeblad* 1950.

Symbol	Numbers
S ₁	0 0 0
----- 3. Division	
S ₂	0 0 1
----- 2. Division	
S ₃	0 1 0
----- 3. Division	
S ₄	0 1 1
----- 1. Division	
S ₅	1 0 0
----- 3. Division	
S ₆	1 0 1
----- 2. Division	
S ₇	1 1 0
----- 3. Division	
S ₈	1 1 1

Figure 3-1 N^{th} -order choices, elements of equal probability

and the signal is a continuously varying electrical function of time.

c) *Hybrid system.* Both discrete and continuous signals are present. An example is pulse-code modulation (PCM).

The individual messages are generated through a series of *selections from a given register*, that is to say, by successive selection from a given collection of symbols of one type or another. For example, a written message is produced by selecting letters from the alphabet of the language in question; speech by selecting particular sounds that the speaker can generate with his or her voice. It is worth noting that the register in general consists of a limited number of symbols: hence we only have a few dozen written characters at our disposal, and speech is physiologically restricted to approximately 50 different sounds. The number of possible messages of finite duration is therefore also limited. The larger the number of different symbols and the longer the duration, the more information the individual message will contain. The amount of information in a single message therefore increases with the number of possibilities.

3 Information in a Discrete System

Written text is a typical example of a discrete message. There are indications that signals in the human nervous system are also discrete, and can therefore be represented by a series of choices.

3.1 Unit of Information

The simplest type of choice is a choice between *two equal possibilities*: 1 or 0, yes or no, heads or tails, or simply, any case of equally possible “either – or” states. We will define the unit of information as the outcome of such a binary *elementary choice*, and we will designate it *1 bit* (abbreviation of “binary digit”). The designation *1 Hartley* has also been suggested, after one of the pioneers of information theory. Further, we will postulate that H *independent* elementary choices provide H bits of information. Based on this, we are able to develop methods of calculating the amount of information in discrete messages.

3.2 N^{th} -order choice

The selection of one symbol from a register of N – that is a group of N possible elements – is called an N^{th} -order choice. To specify how much information such a choice represents, we must reduce it to a series of *independent* elementary choices. The number of these necessary to specify one of the N possible elements is by definition equal to the number of bits of information for the N^{th} -order choice in question.

3.2.1 Choice of N Elements of Equal Probability

We envisage these as N symbols arranged in a column, as shown for $N = 8$ in Figure 3.1. To specify one particular element by means of a series of elementary choices, we proceed as follows: First, we divide them into two equal groups, which represents one elementary choice. Then we divide each group into two sub-groups; two elementary choices will be sufficient to specify one of these. The process of successive division is continued until the desired symbol is isolated from the others. We see that

- 1 elementary choice is needed for $N = 2$
- 2 elementary choices are needed for $N = 4$
- 3 elementary choices are needed for $N = 8$
- etc.

In general,

$$H_N = \log_2 N \text{ bits} \quad (3-1)$$

are required to specify a choice from N equal possibilities.

The groups which arise after each division can be designated by 0 or 1 respectively, as shown in

a priori Probability P_i	Symbol	Numbers	H_i
1/2	S ₁	0	1
----- 1. Division			
1/4	S ₂	1 0	2
----- 2. Division			
1/8	S ₃	1 1 0	3
----- 3. Division			
1/8	S ₄	1 1 1	3

Figure 3-1. The resulting number of digits for each symbol is then equal to H_N , that is, equal to the number of bits necessary to specify the symbols.

The above-mentioned result is strictly speaking only correct if N is a power of 2, so that H_N is an integer. If this is not the case, H_N will be equal to one of the two integers that are closest to $\log_2 N$, depending on which symbol is to be specified.

3.22 Choice of N Elements of Unequal Probability

Each of these possibilities has given (*a priori*) probabilities. Again, we perform successive divisions in elementary choices, that is, into two groups of elements having *equal probability*. The divisions are performed such that the total probability in the groups is equal.

Let us first consider an example. In Figure 3-2, four symbols S_i are given, with probabilities p_i (where $i = 1, \dots, 4$). The divisions are performed as shown, and the groups are designated by 0s and 1s. The number of digits is then equal to the number of elementary choices H_i bits, that are necessary to specify the symbol in question. H_i is different for the different symbols, the infrequent ones representing more information than the frequent ones. We find that

symbol	S_1	S_2	S_3	S_4
H_i	1	2	3	3

The average information per choice is thus

$$H_N = \frac{1}{2} \cdot 1 + \frac{1}{4} \cdot 2 + \frac{1}{8} \cdot 3 + \frac{1}{8} \cdot 3 = \frac{7}{4}$$

Generally, we find that the number of bits necessary to specify a symbol with *a priori* probability p_i

$$H_i = \log_2 \frac{1}{p_i} = -\log_2 p_i$$

Since the sum of the probabilities of all symbols is equal to 1, the average information for the N^{th} -order choice is

$$H_N = \frac{\sum_{i=1}^N p_i \log_2 \frac{1}{p_i}}{\sum_{i=1}^N p_i} = -\sum_{i=1}^N p_i \log_2 p_i \quad (3-2)$$

This expression has a maximum at $p_i = 1/N$, that is to say, in the case of equally likely choices. It is then reduced to Equation (3-1). If the four symbols in the example above had been equally likely, i.e. all $p_i = 1/4$, the information per N^{th} -order choice would have been

$$H_N = \log_2 N = 2$$

which is greater than for any other set of probabilities.

The right-hand side of Equation (3-2) has, except for the sign, the same mathematical form as entropy in thermodynamics for a system whose possible states have the probabilities p_i . It has therefore been designated *negative entropy*.

The fact that the base 2 has been included in the above expressions results from our decision to define the unit of information as a choice between 2 equal possibilities. If we instead had chosen to define the unit as a choice between 10 equal possibilities, we would have obtained common logarithms, etc. Mathematically, there is of course no difference. However, in practical terms, there are good arguments for working with binary numbers. An electrical apparatus, for example, does not need to measure the size of a pulse, but only decide whether the pulse is present or not. A switch, or flip-flop circuit has two

possible states; m such elements therefore have 2^m possibilities, and can therefore store m bits of information. Moreover, from experience, it is natural for a person to divide successively into two equal groups, rather than directly into ten groups, for example.

3.3 Discrete Messages and Signals

A discrete message (signal) will generally consist of a series of symbols $S_1, S_2, \dots, S_i, \dots, S_N$, having the probabilities $p_1, p_2, \dots, p_i, \dots, p_N$, i.e. a series of N^{th} -order choices. Such a process, which depends on a set of *a priori* probabilities, is called a *stochastic* process. It is characterised by certain statistical properties, which are determined by these probabilities. An example of such a discrete message is ordinary text, in which the symbols are selected from the alphabet of the language in question, which has known letter frequencies of occurrence, etc.

In the following treatment, we will assume that the messages (signals) we are concerned with are *statistically stationary*, that is to say that their statistical characteristics are invariant towards translation in time of the relevant time series. Stated in another way: we assume that the ergodic hypothesis is valid. The statistical characteristics can then be determined as time averages for a single message of sufficient duration, or as the average for all messages in the same class at a given point in time.

3.31 The Symbols are Independent

If the individual messages consist of one symbol that can be selected from N possible symbols, there are N possibilities. If they consist of two symbols, there are N^2 possibilities, and so on. The total number of messages possible with n such N^{th} -order choices is $M = N^n$. If these are equally possible then the average information per message (consisting of n symbols) is therefore

$$H_M = \log_2 M = n \log_2 N$$

and the average information per symbol

$$H_S = \log_2 N \quad (3-3)$$

If the choices do not have equal probability, the average information per symbol according to the ergodic hypothesis

$$H_S = - \sum_{i=1}^N p_i \log_2 p_i \quad (3-4)$$

It can be shown that this expression has a maximum value at $p_i = 1/N$, in other words, when all symbols are equally possible.

H_S is zero only when all values of p_i are zero except one, and this equals 1, i.e. $N = 1$. The message consists in this case of a series of one

single symbol, and is therefore characterised by an a priori probability of 1 (i.e. certainty), and contains no information.

3.32 The Symbols are Dependent

In practice, the symbols are often dependent. Thus, in Norwegian *e* is very often followed by *n* or *r* in word-endings, and there is therefore a greater conditional probability of *n* or *r* occurring after *e* than, for example, after *l*.

It can be shown that Equation (3-4) is valid also in the case of dependent choices, if we for p_i introduce the *conditional* probability for symbol S_i , based on knowledge of all previous choices.

With dependent choices, the results of Section 3.31 can be considered as a first approximation. They will give too high values of H_S , because the uncertainty of each choice is smaller when those choices are dependent.

3.4 Redundancy

We define the *redundancy* of an information source as

$$R = \frac{H_{S_{\max}} - H_S}{H_{S_{\max}}}$$

where $H_{S_{\max}}$ is the maximum negative entropy we could have obtained for the same symbols (if these were independent and equally probable). The redundancy is an expression of the correlation in the message.

For example, let us consider the English language. There are certain a priori conditional probabilities for the occurrence of digrams (two consecutive letters), trigrams (three consecutive letters), etc, and for certain words. The redundancy for everyday English is found to be approximately 50 %, if we do not consider statistical structures over more than eight letters. This means that half of written English is determined by the structure of the language, while the remainder is chosen freely. The above-mentioned figures have been determined in different ways: by calculating the negative entropy of statistical approximations to English, by eliminating a certain proportion of the letters in a text and letting another person read it, and by cryptographic methods. Here, it is worth mentioning that there is a considerable difference between the redundancy of Shakespeare's English with a large vocabulary and that of so-called "basic English" with a vocabulary of about 850 words. In the first case, it is possible to express oneself briefly and concisely, and the redundancy is low. In the second case, a large number of words is needed for an exact description, and the redundancy is therefore high. Similar considerations can be made for the different forms of the Norwegian language.

Another example is an address label on a postal package recently received by the author. It read
 Forsvarets Forskningsinstitutt,
 [correct: Forsvarets Forskningsinstitutt]
 Avd. for Radar
 Bergen ... Norway

Thanks to the redundancy in this message, the package reached its destination.

Also in television signals, redundancy is common; hence the background of the picture can remain unchanged for long periods, while a person in the foreground makes small movements.

In an efficient communications system, we will often remove part of the redundancy at the input to the channel. The more we remove, the more important the remaining symbols become. How much we are to remove therefore depends on the impairments (noise) on the channel.

3.5 Coding

The discrete messages consist of symbols of varying frequency of occurrence, and can therefore be represented by a series of N^{th} -order choices (which as a rule have unequal probability). Each of these can be reduced to a series of elementary choices; this process is called binary coding. Theoretically, optimal binary coding requires an average of H_S elementary choices/symbol, where H_S is given by Equation (3-4).

The example in Section 3.22 illustrates a very simple coding process. The binary codes are identical to the numbers that arise from the successive divisions into two groups of equal probability. It follows from this that the most frequently occurring symbol has the shortest code, and that the code becomes longer the more infrequent the symbols are.

We can think of coding as a "stretching" of the time scale according to the probability of occurrence for the individual symbols, that is, a *statistical adaptation* of the channel to the class of message that it is to transmit. Of course, this requires a time delay, which in the case of optimal coding, can be infinitely long. H_S in Equation (3-4) is therefore to be considered as a lower bound in practice. It is, however, generally easy to calculate how close we are to this optimal state and thus indicate the degree of efficiency of the coding. This will depend on the memory of the transmitter.

Of course, we do not need to use binary coding: the Morse code, for example, is quaternary. For English and Norwegian text, this is certainly not bad, as the most commonly occurring letters such as *e* and *t* have the shortest codes, and so

on. For the Czech language, which has completely different letter frequencies, it is probably less efficient.

4 Information in Continuous and Hybrid Systems

The information content in *continuous* messages and signals can be determined by reducing them to discrete signals. This is done in two stages: *sampling* of the signals gives a series of ordinates, which can then be *quantized*. We shall now look at these two processes in more detail.

4.1 Sampling

The function of time $f(t)$ is assumed to contain only frequency components in the bandwidth $0 - W$ p/s. Intuitively, we know that $f(t)$ cannot then change to a substantially new value in an interval of time $\frac{1}{2W}$, which is half the period of the highest frequency. On this basis, pulse modulation was introduced, and experiments were performed to examine the intelligibility of signals with sampling frequencies from 1 up to 3 times W . It was found that a factor of around 2 was sufficient to achieve good quality. We will now demonstrate that $f(t)$ is defined exactly by sampling at a frequency of $2W$ per second.

$f(t)$ is multiplied by a function $S(t)$, consisting of delta pulses repeated at a frequency f_r . The product $f(t) \cdot S(t)$ is then a series of ordinates of $f(t)$ separated by a distance $1/f_r$. $S(t)$ can be broken down into a Fourier series consisting of a DC component, the fundamental frequency and the higher harmonics.

As shown by the theory of pulse modulation, multiplication by $f(t)$ will result in sidebands around each of these components with the same shape as the spectrum of $f(t)$, that is with a width of W . This is illustrated in Figure 4-1. If we are to separate the spectrum for $f(t)$ around the DC component from the lower sideband around the fundamental, the two must not be mixed. The condition for this is

$$f_r - W > W$$

or

$$f_r \geq 2W$$

In most practical cases, it is simplest to sample periodically, but this is not necessary. We can, for example, also determine $f(t)$ uniquely by means of ordinates with varying time intervals, or by means of ordinates and derivatives, as long as the frequency of sampling is at least $2W$ per second.

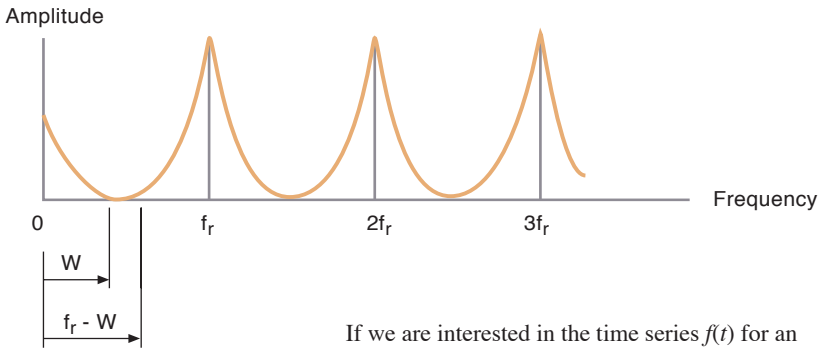


Figure 4-1 Frequency components for $S(t)$, with associated sidebands for $f(t)$

If we are interested in the time series $f(t)$ for an interval of T , then

$$n = 2TW \text{ samples} \quad (4-1)$$

will be sufficient to characterize $f(t)$ uniquely and exactly.

4.2 Quantization

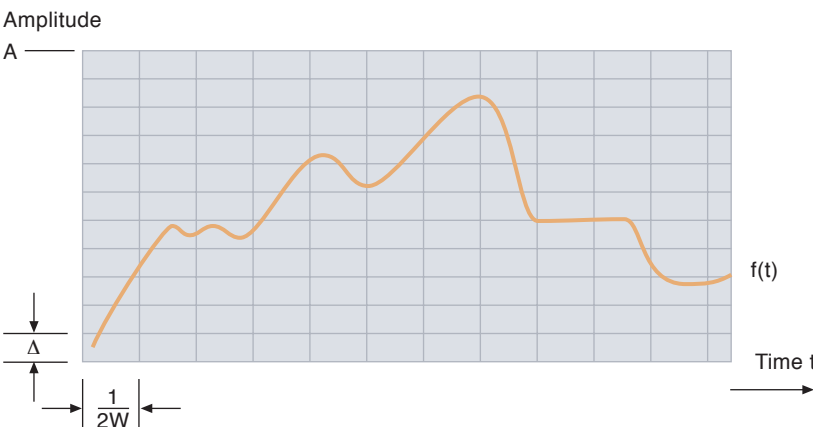
Through sampling, we have reduced a continuous function of time $f(t)$ with limited bandwidth to a series of ordinates which exactly define $f(t)$. These can nevertheless vary continuously in amplitude, and hence an infinite number of elementary choices, i.e. an infinite number of bits of information, will be needed to specify any one of them. However, there is no need to specify the amplitudes with greater accuracy than a certain *tolerance* Δ , which is primarily determined by the level of noise. We will therefore only specify certain amplitude steps, and we call this quantization. If the maximum value of the amplitude of the function is A , we will be able to distinguish between a total of

$$N = \frac{A + \Delta}{\Delta}$$

We wish to express N in terms of a signal power P_s , and a noise power P_n , which are the characteristic expressions for statistical messages and noise. Knowing the amplitude distribution of fluctuation noise, we can, on the basis of experience, set

$$N = \frac{A + \Delta}{\Delta} = \sqrt{\frac{P_s + P_n}{P_n}} \quad (4-2)$$

Figure 5-1 Continuous signal



5 The Information Capacity of the Channel

5.1 Derivation and Definition

We will now determine how many bits H of information it is possible to transmit in time T through a channel with bandwidth W and signal-to-noise ratio $K_s = P_s / P_n$ (reduced to receiver input).

Figure 5.1 shows a continuous signal $f(t)$ in a grid where the division along the time axis is the sampling interval $\frac{1}{2W}$ and the division along the amplitude axis is the tolerance Δ .

According to Equation (4-2), there are

$$N = \sqrt{\frac{P_s + P_n}{P_n}} = \sqrt{1 + K_s}$$

possible levels in each time interval, and for a signal of duration T there are, according to Equation (4-1),

$$n = 2WT$$

such time intervals.

In a single time interval there are N possible amplitudes, after two time intervals there are therefore N^2 possibilities, and so on. The total number of possible signals is therefore

$$N^n = \sqrt{1 + K_s}^{2WT} = (1 + K_s)^{WT}$$

In other words, there is a limit to the number of signals we can transmit through this type of channel. If they are equally probable, the amount of information per signal in bits is maximum, that is

$$H = \log_2 (1 + K_s)^{WT} = WT \log_2 (1 + K_s) \quad (5-1)$$

The information the channel can transmit per time interval is therefore

$$C = H/T = W \log_2 (1 + K_s) \quad (5-2)$$

This quantity is called the *information capacity* of the channel, and its units are bits/sec.

If we are to exploit this capacity fully to transmit a message containing an amount of information H in a time T , *optimal coding* will be required (see Section 3.5), that is, the transmitter must convert the message to a signal which is completely statistically adapted to the channel. We will call such a communication system an *ideal* system. It is *not possible* to transmit more information per time interval than C .

5.2 Discussion of an Ideal System

Equation (5-2) applies to all values of K_s . For $K_s \ll 1$, $\log_2(1 + K_s) \approx K_s \log_2 e$, i.e.

$$C \approx 1.443WK_s$$

For $K_s \gg 1$

$$C \approx W \log_2 K$$

Equation (5-2) shows the relationship between the parameters W , K_s and C . For the same information capacity C in two cases, marked 1 and 2,

$$(1 + K_{s1})^{W1} = (1 + K_{s2})^{W2}$$

$$1 + K_{s1} = (1 + K_{s2})^{W2/W1}$$

or, approximated for large values of K_s

$$K_{s1} \approx K_{s2}^{W2/W1}$$

We can thus reduce the bandwidth W if we simply increase the signal-to-noise ratio enough. A reduction in W by half will therefore require an increase in K_s to the second power. Conversely, we can increase W and thus manage with a substantially lower K_s .

The reduction in bandwidth could, for example, be achieved in the following manner: instead of specifying each sample (at intervals of $\frac{1}{2W}$), every second sample and its derivative are transmitted as a composite number. In this way, the necessary bandwidth can be halved. On the other hand, the signal-to-noise ratio must be raised to the power of 2 to enable us to read the composite number with the same accuracy.

It should be mentioned that the bandwidth enters into the expressions for $K_s = P_s / P_n$, since, for fluctuation noise,

$$P_n = W \cdot 4kT_{\text{abs}}$$

where k is Boltzmann's constant and T_{abs} = temperature.

Equation (5-2) can be derived exactly by means of multi-dimensional geometry.³⁾ This kind of treatment will also lead to more wide-reaching results, such as an explanation of the threshold problem in broadband modulation systems.

6 Applications

6.1 Miscellaneous

There are reasons to believe that the nervous systems of living organisms consist of nerve cells with two possible responses: 0 or 1, exactly like a switch. The signals can therefore be characterised by a series of binary choices, and the

concepts and results of information theory open up a possibility of studying the transmission of information in *nervous systems*.⁶⁾

The human central nervous system contains something of the order of 10,000 *million* cells. The most complicated computer so far built, ENIAC, has 10,000 binary elements, which is about the same number as in the nervous system of an earthworm. The large number of cells in the human brain makes it possible for a signal to be transmitted by several parallel routes, so that a fault in a single cell in a chain does not significantly interfere with the transmission. In an electronic computer, however, such a fault would usually lead to a completely erroneous result in all subsequent calculations, because the signal usually can only travel by one route. The results of information theory have become important in connection with the operation and correction of errors in computers.⁷⁾

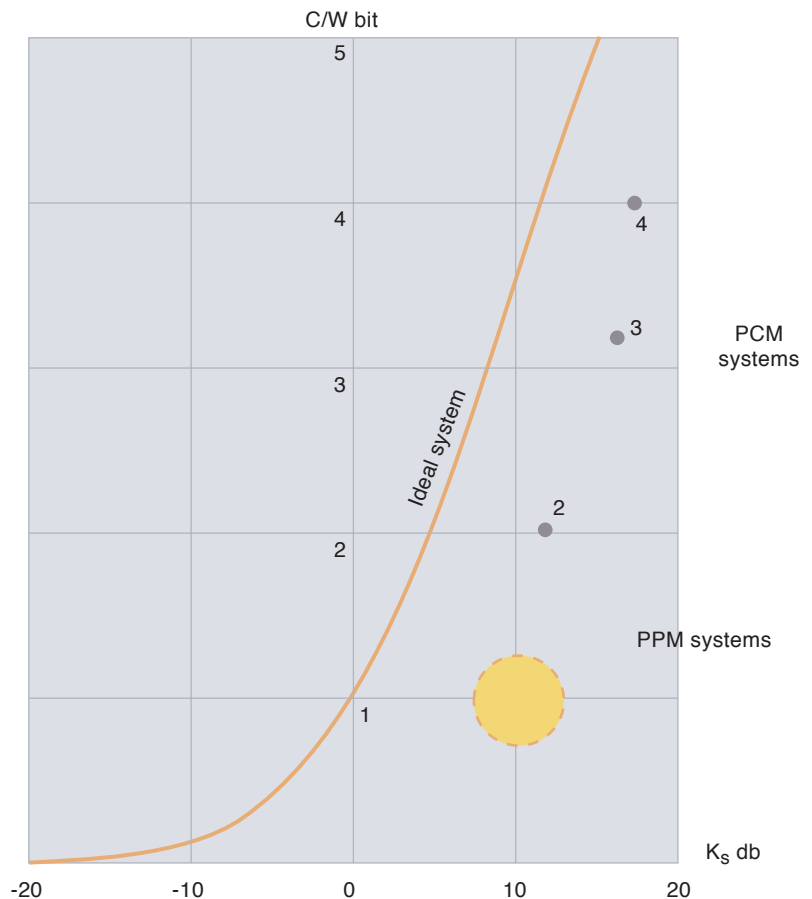
Moreover, a complete theory for transmitting secret information (*cryptology*) has been developed.⁸⁾

6.2 Practical Communications Systems

The equation

$$C / W = \log_2(1 + K_s)$$

Figure 6-1 Graphic representation of $C/W = \log_2(1 + K_s)$. Some typical practical systems are indicated



is represented graphically in Figure 6-1. The curve shows how much information it is generally possible to transmit per unit time and per cycle of bandwidth for a signal-to-noise ratio K_s . In general, the ideal values will not be attainable in practice, since this would require optimal coding, which can result in an infinitely long time delay. However, it is possible to calculate how far a practical system will be from the optimal system; hence the diagram indicates points for typical pulse position modulation (PPM) systems and binary, tertiary, and so on, pulse code modulation (PCM) systems, without delay in the transmitter.

We will define the information efficiency η as the ratio of the amount of information per second that is actually transmitted to the maximum amount of information that could have been transmitted through a corresponding ideal system. To put it another way

$$\eta = \frac{\text{inf./sec. in received message}}{\text{information capacity}} = \frac{W_m \log_2(1 + K_m)}{W_s \log_2(1 + K_s)} \quad (6-1)$$

where the indices m and s represent the message and the signal, respectively.

Let us consider the information efficiency for a number of practical modulation systems.⁵⁾

Amplitude modulation (AM):

single sideband $W_m = W_s$ and $K_m \approx K_s$
or $\eta \approx 1 = 100\%$
double sideband $W_m = 0.5W_s$ and $K_m \approx K_s$
or $\eta \approx 0.5 = 50\%$

Frequency modulation (FM):

η is in the region of 20 – 2 % when the modulation index μ (= frequency deviation/ W_m) varies between 1 and 100. η decreases as μ increases; when the modulation index increases, i.e. increase of the bandwidth W_s , a better message-to-noise ratio is obtained, but the information efficiency decreases. The reason for this is that the denominator increases with the bandwidth W_s , while the numerator only increases with the log of K_m .

Pulse modulation:

With the exception of PCM, the different systems have an information efficiency of the same order of magnitude as FM. This can be explained in the same way, namely that the bandwidth W_s becomes significantly larger than W_m , so as to obtain a desired message-to-noise ratio for a relatively low K_s .

Pulse code modulation:

Here, η is around 40 %. As is known, PCM is the only practical system devised so far for which K_m increases with bandwidth W_s . An increase in W_s will therefore not reduce η , as with other types of pulse modulation.

Previously, it has only been possible to determine a communications system's efficiency in transmitting information by means of *comprehensibility tests*. This is a subjective measure, unless we allow a very large number of people to act as information source or destination, and use statistical methods to find average scores. This would, however, prove expensive, because the system must first be built and then subjected to time-consuming tests. However, on the basis of the general definition Equation (6-1) for information efficiency η , we are now able to calculate (that is, *objectively* assess) the ability of an individual communications system to transmit information of various types.

Literature

- 1 Fano, R M. *The Transmission of Information*. Massachusetts Institute of Technology, Research Laboratory of Electronics. (Technical Report No. 65.)
- 2 Shannon, C E. A Mathematical Theory of Communication. *Bell Syst. Tech. J.*, 27, 623–656, 1948.
- 3 Shannon, C E. Communication in the Presence of Noise. *Proc IRE*, 37, 10–21, 1949.
- 4 Tuller, W G. Theoretical Limitations on the Rate of Transmission of Information. *Proc IRE*, 37, 468–478, 1949.
- 5 Clavier, A G. Evaluation of Transmission Efficiency According to Hartley's Expression of Information Content. *Elec. Com.*, 25, 414–420, 1948.
- 6 Wiener, N. *Cybernetics*. Wiley, 1948.
- 7 Hamming, R W. Error Detecting and Correcting Codes. *Bell Syst. Tech. J.*, 29, 147–160, 1950.
- 8 Shannon, C E. Communication Theory of Secrecy Systems. *Bell Syst. Tech. J.*, 28, 656–715, 1949.

Excerpts from the Discussion After the Lecture

Garwick:

The lecture was very interesting. I have noticed a couple of points that need further clarification.

1. The lecturer mentioned that when a time series is to be transmitted without distortion by means of sampling, it is sufficient to sample with a time interval equal to or greater than $\frac{1}{2W}$, where W is the bandwidth. But what is the bandwidth? In the case of a completely square spectrum, this is simple enough, and similarly if the spectrum has some other shape, but is limited such that frequency components above a certain limit do not exist. In practice, the bandwidth will often have infinite extent. The frequency spectrum's sidebands will also have infinite extent, and the proof will therefore not be mathematically valid.

2. During his treatment of quantising, the lecturer stated that:

$$N = \frac{A + \Delta}{\Delta} = \sqrt{\frac{P_s + P_n}{P_n}}$$

The left-hand side is correct, but it is not immediately obvious that the right-hand side is equal to the left-hand side, since it is the amplitudes that characterise the noise, and not the power.

Knudtzon:

1. It was assumed that the message was limited in frequency to the bandwidth $0 - W$ p/s. Whatever the shape of the spectrum within this bandwidth, the bandwidth is therefore W . It is easy to demonstrate mathematically that an infinitely long message can have a limited frequency spectrum. If the message has passed through an *ideal* low-pass filter whose upper frequency limit is W , the bandwidth will be W , irrespective of the behaviour of the system function in the pass band. In practice, the attenuation in the rejected band can be made large, but not infinite, for all frequencies greater than W . We will therefore introduce a tolerance in the same way as in quantising, and the bandwidth will be that frequency at which the attenuation definitively exceeds the tolerance.

2. I stressed that from our knowledge of the amplitude distribution of fluctuation noise, the two expressions can be equated to each other. In fact, it is precisely the *power* that characterises the noise. A mathematical treatment based on multi-dimensional geometry has been presented by Shannon.*)

Gaudernack:

The lecturer has pointed out that a piece of information is not completely known until one has full knowledge of its statistical properties, and that this knowledge is necessary to be able to

dimension the network statistically correctly. What significance does this have for the practising engineer who wishes to try out using these new methods? What investigations must be carried out before it is possible to say that one knows a class of information completely?

Knudtzon:

I will mention some examples. In telegraphy, it is possible to set up tables showing the frequency of occurrence of letters, digraphs, tri-graphs and words. Such tables exist for the English, French, German, Italian and Spanish languages at least. In the case of telephony, one needs the correlation function, which can be measured electronically. The result depends to a large extent upon what one wishes to transmit, for example to what extent the character (intonation, emotion, atmosphere) is to be preserved. These matters are being subjected to extensive measurements at the Psycho-Acoustic Laboratory at Harvard, among other places. Also television signals are being studied: it should, for example, be unnecessary to transmit a stationary background in a television picture continuously.

I will briefly summarise the technique for electronic determination of the correlation function. The autocorrelation function for a time function $f(t)$ is defined as:

$$\varphi(\tau) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^{+T} f(t)f(t + \tau)dt$$

The function $f(t)$ is sampled periodically at a frequency of f_s in pairs separated by intervals of τ_k seconds, giving us the samples $\alpha_1, \beta_1, \dots, \alpha_n, \beta_n, \dots$. Pulses are generated with heights proportional to α_n , and widths proportional to β_n , such that the areas represent the products $\alpha_n\beta_n$. The average is then taken of the integrated product over the observation time T , and we get the following expression for the autocorrelation for τ_k

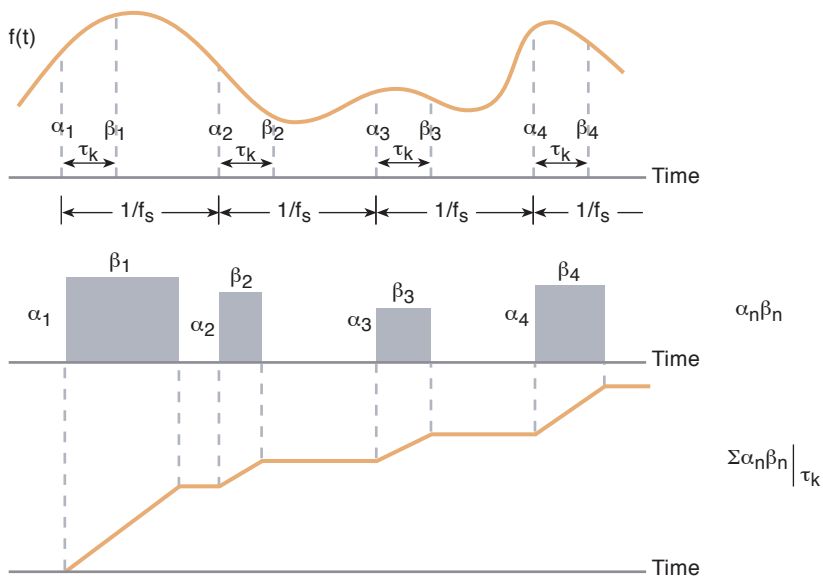
$$\varphi(\tau_k) \approx \frac{1}{Tf_s} \sum_{n=1}^{Tf_s} \alpha_n\beta_n | \tau_k$$

Then, the measurement is carried out for τ_{k+1} , and so on. Correlators of this type have been built at the Research Laboratory of Electronics at MIT.

Falnes:

The lecturer has treated statistically distributed noise. In practice, impulsive noise is also experienced. Although the average noise voltage is low, strong spikes can completely obliterate a message such as a telegraphy character.

*) Proc. IRE 1949 pp 10-21.



The Hell System which has replaced ordinary telegraphy is less sensitive to impulsive noise, because each character is represented by a larger number of impulses. As the lecturer pointed out, one therefore achieves better signal-to-noise ratio, at the expense of bandwidth.

Garwick:

I would like to mention some examples which illustrate how the transmission of unnecessary data provides a check of the accuracy of the transmitted signal. First, let us consider how the number 25 can be written in the binary system.

The number is written as 11001. This is the minimum number of necessary characters. If one of them is reproduced incorrectly, the number itself becomes incorrect. In a coded decimal system, the number can be written as 0010, 0101. The first group represents the figure 2, the second group the figure 5. In each group, there are some character combinations which do not represent a number. Thus, if such an impossible combination is received, this indicates that an error has occurred. Even safer systems can be achieved with other methods of coding. This method is used, for example, in a mathematical computer constructed at the Norwegian Defence Research Centre.

Helmer Dahl:

It might be interesting to stress the importance of the derived equation for information capacity from a practical point of view. The equation reads:

$$C = \frac{H}{T} = W \log_2(1 + K_s)$$

and represents the *maximum* amount of information that can possibly be transmitted in a given time interval, at a bandwidth W and signal-to-

noise ratio K_s . The equation defines in a way the boundary between what is possible and what is impossible. For example, it can prevent us from trying to transmit a message through a channel that is not capable of carrying that message.

On the other hand, we can only achieve the maximum information capacity by using optimal coding, and this can demand expensive and complex technical equipment. Hence, it may well be economically justifiable to use a more primitive type of coding, and therefore a greater bandwidth and output power than is strictly necessary, technically speaking. This is especially the case when one attempts to save power by using large bandwidths, since one then makes use of far greater bandwidths than the information capacity actually requires, in order to obtain simple coders and decoders (modulators and demodulators). This may also have historical causes, since up to now modulation systems have been developed without considering the concept of information capacity, and it is possible that by using information theory, more suitable methods may be found which are both economically and technically efficient.

Nevertheless, there is reason to believe that maximum information capacity and optimum economics only rarely will coincide. In cases where bandwidth is cheap, such as in UHF systems, it may be worthwhile to use large bandwidths in order to save power by simple technical means. On the other hand, it is interesting that in cases where bandwidth is expensive, as in cable communications, single sideband transmission has proved to be a suitable system which approximates closely the theoretical optimum for transmission. Carson demonstrated the special characteristics of this system long ago, on the basis of other criteria. Based on information theory, we can see his results in a broader perspective, as one of many possible optimal solutions.

Knudtzon:

Finally, I would like to mention that in a communications system containing noise it is possible to calculate various negative entropies: at the input to the channel, at the output from the channel, and corresponding expressions for the conditional probabilities between input and output. Many important conclusions can be derived on the basis of these calculations. It can also be demonstrated mathematically that the negative entropy will always decrease on passing through the system, and the average information can therefore never increase in an isolated system. This is completely analogous to the second law of thermodynamics.

Statistical Communication Theory

A Brief Outline of the Problem¹⁾

Graduate Engineer *Nic. Knudtzon*, M.N.I.F.
Norwegian Defence Research Establishment, Bergen



Dr. Nic. Knudtzon (80) obtained his Engineering degree from the Technical University of Norway, Trondheim in 1947 and his Doctor's degree from the Technical University in Delft, the Netherlands in 1957. 1948–1949 he was with the Research Laboratory of Electronics, Massachusetts Institute of Technology, working with information theory and experiments. 1950–1955 he was with the Norwegian Defence Research Establishment, Bergen, working on the development of microwave radio links; and from 1955 to 1967 he was Head of the Communications Division at Shape Technical Center in The Hague, Netherlands, where his efforts went into the planning of military telecommunications networks and systems in Western Europe. From 1968 to 1992 he was Director of Research at the Norwegian Telecommunications Administration, working on the planning of future telecommunications systems, networks and services. Dr. Knudtzon has been member of government commissions and various committees, including the Norwegian Research Council, the National Council for Research Libraries, the International Telecommunications Union, EURESCOM, etc.

This is a translation into English of the paper "Statistisk kommunikasjonsteori", which appeared in *Teknisk Ukeblad* on November 16, pp. 883–887, 1950. The translation was done by Berlitz GlobalNET and final quality control was done by Geir R. Øien.

The schematic diagram in Figure 1 illustrates a communication system, which consists of an information source, channel and destination, with the channel consisting of sender, transmission and receiver. The information source generates *messages*, which can either be discrete²⁾, as used in telegraphy, or continuous as in telephony. The sender converts the message to a *signal* that suits the connection, which is a cable or radio connection. The receiver attempts to retrieve the original message by carrying out the inverse process of the sender. The destination is the person or device that the message is bound for.

Three things characterise a communication system of this type:

- It is to transmit information.
- The equipment and channel have a limited frequency band.
- Noise occurs, while at the same time signal strength is lost. A certain signal/noise relationship is therefore required in order to achieve a specified message/noise relationship at the destination.

These facts may seem rather obvious to us today. However, it is only in recent years that we have fully taken the consequences of them, and in so doing have arrived at a general communication theory. In order to see this more clearly we will briefly consider some fundamental stages in the development of communication techniques.

The original idea was that the electrical function of time representing the message at receiver-out, should be an exact reproduction of the time series for the message at sender-in. Consequently, what was desired was the elimination of the distance between the information source and destination, hence the prefix "tele" (far off) in telegraphy, telephone, television, telemetry etc. It was expected that part of the signal output would be lost in the channel, but that it had a *frequency-dependent response* just like all other electrical networks was a surprise to most of the communication engineers of the time. At that time only slow telegraphy and telephony had been developed and the systems' frequency characteristics were detailed based on practical tests, most of which were comprehensibility tests. Due to the properties of these special types of signals and the human ear, only the amplitude response was taken into consideration, while the phase response was almost completely ignored. Unfortunately, this became customary and is often still the case today. This further led to electrical networks being studied exclusively in the *frequency range* for many years, by means of so-called "sinusoidal analysis" (steady state analysis). The network is then uniquely determined by its system function $H(j\omega)$, which can be defined as the relationship between a signal's frequency spectra at output and input. This is generally a complex-valued function.

The next milestone of significance to our discussion was the discovery of the radio valve, which

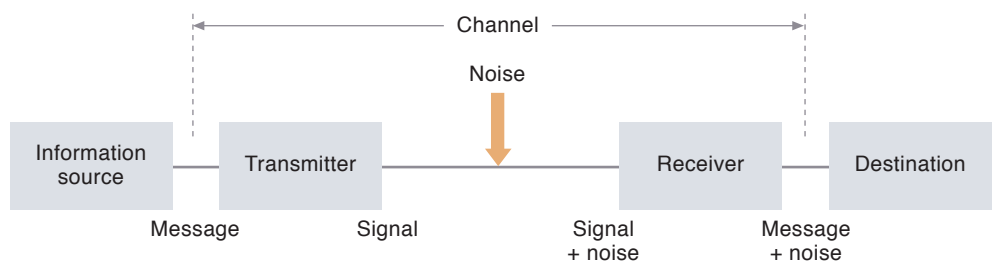
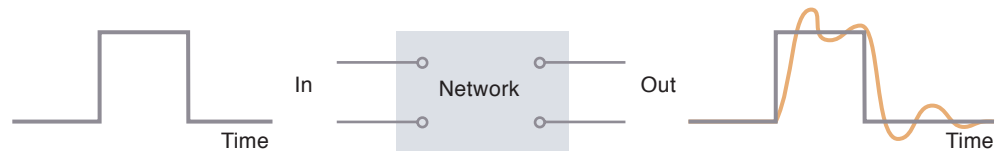


Figure 1 Communication system

¹⁾ Lecture presented at a study session for radio technology and electro-acoustics at Farris Bad, 16–18 June 1950.

²⁾ Non-continuous.

Figure 2 Typical deformation of pulse in an electrical network



enabled the amplification of the signals. The losses could now be gained back, and this was of importance for transmission over both cables and radio. However, a new problem arose as different types of *noise* manifested themselves. Fluctuation noise that enters via the aerial and is generated in valves and resistors, occurs at all frequencies; the power is proportional to the bandwidth. In order to combat the noise, other *modulation methods* were introduced in addition to the amplitude modulation (AM), which was known from multiplexing channels, amongst other things. Common to all modulation methods is that a low frequency message can be transmitted as a high frequency signal. Frequency modulation (FM) was the first solution. This idea was first put forward because it was hoped that one could manage with narrower frequency bands than with AM, so that the effect of the noise would be reduced. However, the opposite proved to be true, as practical tests showed that the broader the frequency band became, the more the message/noise relationship improved for the same transmit power. FM is thus an illustrative example of how fumbling progress has been made. After further experimentation, the different pulse modulation systems were developed one by one; as we know, intense work is currently being carried out on pulse code modulation (PCM). Detailed investigations of all these systems show that for the same message/noise relationship at receiver-out, the sender's average signal power P_S can be reduced when the bandwidth W is increased, if only the signal/noise relationship $K_S = P_S / P_N$ (reduced to sender-out) is higher than a certain threshold value. Thus, *noise reduction requires a broad frequency band*. It should be added that when calculating the message/noise relationship for the different modulation systems, only messages that are sinusoidal functions of time have been so far considered.

As is known, during pulse modulation, a series of the message's ordinates are transmitted instead of the entire continuous time series. This process is known as *sampling*. It can be shown³⁾ that the message is completely and exactly determined by the samples, if these are taken more often than the equivalent of twice the highest

frequency component. The samples are *quantized* in PCM, which means that only certain discrete amplitudes are transmitted. The stages between these are determined by the distortion that is thereby allowed. By sampling and quantizing, we have managed to *convert the continuous message to a discrete message*.

Television was developed around the same time as the modulation methods. The television signals consist of a series of pulses; a typical example of how such a pulse is deformed in an electrical network is shown in Figure 2. It is extremely important that the pulses do not overlap. The development of the television therefore led to studies in the networks' responses in the *time domain* for a pulse input of a suitable form. It has proved to be beneficial in calculations to use the so-called δ pulse, which is defined as follows:

$$\delta(t) = 0 \text{ for } t \neq 0$$

$$\int_{-\infty}^{+\infty} \delta(t) dt = 1$$

All other time series can be thought of as composed of such δ pulses. The network is subsequently precisely determined by its δ pulse response.

It may be appropriate here to explain the connection between the system function $H(j\omega)$ and the δ pulse response $h(t)$ for an electrical network. This has been done in detail in the Appendix. The results show that "sinusoidal analysis" and " δ pulse analysis" are closely connected and can be derived from each other. If we choose to work in the frequency range, it has to be stressed that the amplitude response and the phase response are equally important, and that they cannot be specified independently of each other.

Servo mechanisms were particularly developed during and after World War II, as there was a great need for control equipment for artillery etc. The reason why development in this area has happened so quickly, is that it quickly became apparent that the problems had much in common with those that were previously

³⁾ "Information Theory" lecture presented at a study session for radio technology and electro acoustics at Farris Bad 15–18 June 1950. ETT no. 30, 1950.

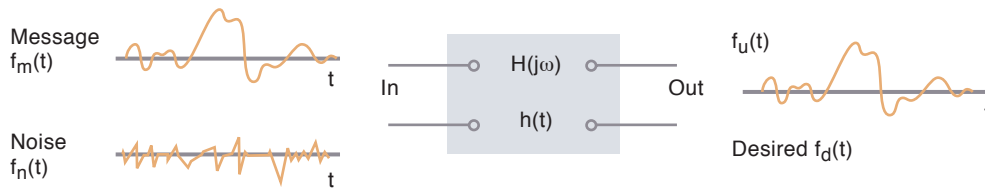


Figure 3 The filter problem

encountered in the network theory, and thus known results could be utilised. However, the opposite effect has also taken place. An important problem that has had major consequences occurred in connection with anti-aircraft guns, where the target can move quickly and in irregular paths. It was therefore important to be able to predict the position of the target for the length of time that it took for the projectile to reach it. In order to solve this problem a thorough study had to be made of the aircraft's path as a function of time. A pilot under fire will naturally steer in curved paths, but the freedom of movement is limited by the construction of the aircraft. Thus, for example the radius of curvature cannot be made smaller than a certain size depending on the aircraft's design, weight and speed. The flight path cannot therefore be regarded as known in advance, but the path's *statistical properties* can be determined for the types of aircraft that are targeted. Consequently, this is what needs to be borne in mind when constructing the predictor.

We are faced with exactly the same problems in communication engineering. The individual messages that we want to transmit over the system cannot be known in advance if they are to contain any information for the destination. There is therefore no point in transmitting a sinus function or a single δ pulse, as their time response is determined for all time. The message received could therefore be predicted immediately and the easiest thing would be not to send it at all. "Sinus analysis" and " δ pulse analysis" have subsequently no real justification. On the other hand, there is reason to emphasise that only a *limited number* of signals can be transmitted over a channel with a certain frequency band and signal-to-noise ratio. The communication system is therefore constructed for a *class of messages with certain known statistical characteristics*. For example, these could be amplitude distribution, correlation functions, power density spectra, letter and word frequencies etc. We will assume that the messages are *statistically stationary*, i.e. that these characteristics are invariant with respect to a shift of the time scale. This is the nature of fluctuation noise.

Let me now consider the *filter problem* illustrated in Figure 3. A message $f_m(t)$ and noise $f_n(t)$ are input to a linear electrical network, which we wish to construct so that the time function $f_u(t)$ is as an exact reproduction of message-in $f_m(t)$ as possible. Completely exact rendering is out of the question, due to stray capacitance etc., which limits the frequency band, and if this was not limited, we would get infinitely strong noise. The best option is therefore to construct the system so that we get the least possible error between the actual time series $f_u(t)$ and the desired time series $f_d(t)$, which in this case with the filter is identical to $f_m(t)$. We must first decide what we mean by error. It is natural here that we specify this in the time domain. If we directly let $[f_u(t) - f_d(t)]$, the error will be a function of time. If we take the mean over a long time, we see that positive and negative failures will cancel each other out and this is not what we want. We bypass this problem by taking the mean of the square of the difference. This means that the system will be determined by

$$\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^{+T} [f_u(t) - f_d(t)]^2 dt = \text{minimum}$$

This error criterion is physically sound for many applications and leads to a mathematically solvable problem. We will call the corresponding network *statistically optimal*. When solving the problem mathematically, it is shown⁴⁾ that the correlation functions for the classes of time series that occur at input and output, determine the network's system function. The network will always be realisable and stable, as the requirement for this is taken into consideration in the solution. It is also possible to derive an expression for the size of the error. Above we have considered the filter problem where $f_d(t) = f_m(t)$. Generally, however, we can specify $f_d(t)$ as we wish, and decide the relevant network in a similar way as for the filter. For a statistically optimum differentiator we thus set $f_d(t) = f_m'(t)$, and for a statistically optimum predictor $f_d(t) = f_m(t + \alpha)$, where α is the desired prediction time. It is worth noting that we can thus, without difficulty, demand a *shift* α in the time domain. The prob-

⁴⁾ "Statistically Optimal Networks", lecture presented at a study session for radio technology and electro-acoustics at Farris Bad, 15–18 June 1950. ETT no. 30, 1950.

lem of constructing statistically optimum networks for classes of time functions with known correlation functions is therefore solved in theory, and practical applications have also been found. It is apparent that the problem is posed very differently and is based on a much sounder foundation than previously known filter theory.

It is important to have a clear view of what we really want to transmit. We shall use telephony as an example. Since the ear does not react to smaller distortions of the phase, an exact reproduction of the microphone current is not essential. We know from experience that a frequency band of at least 4,000 p/s is required to transmit speech of a reasonable quality, so that the user understands what the information source is relaying, and also to some degree gets an impression of its emotions and feelings as is the case in direct speech. Since a very large part of the conversations over the commercial telephone network convey emotions, emphasis here is put on the transmitted message giving as close to the same impression to the listener as if the information source was addressing him directly. Practice has shown that the price of a telephone channel increases proportionately with the bandwidth, and for financial reasons it is therefore desirable to reduce this. This can be done in a number of ways; we shall briefly discuss some of these below.

- a) Analysis-synthesis telephony. The microphone current is analysed in the sender, codes are transmitted for the tone type and power spectrum's form, the codes then influence an "artificial voice" (oscillator set) in the receiver. The necessary bandwidth for the same comprehensibility as that of an ordinary channel is approximately 400 p/s, i.e. reduction with a factor of 10. Systems of this type were built before the beginning of the 1940s by Bell Telephone Laboratories. They are known as vocoders.
- b) Sound code telephony. Speech is physiologically limited to approximately 50 different sounds, and someone that talks quickly can barely express more than 5–6 per second. If every sound is given a code, a frequency band of approximately 40 p/s should be sufficient.
- c) Sound group telephony. If we are only interested in transmitting certain sound groups (special sound combinations or words), the bandwidth for special systems can conceivably be further reduced by a factor of 10.

If any of these methods of bandwidth reduction are to be used, the terminal equipment will of course become more expensive so that they can only be beneficial for long transmissions. How-

ever, we are here primarily interested in the fact that it is possible (without time delay) to transmit *comprehensible* speech over considerably narrower bands than those that are currently in use. This is done at the expense of the different information sources' individual characters; with drastic bandwidth reduction the listener will understand well enough what is being said, but will not be able to determine whether the information source is male or female or happy or sad. In many cases, for example in systems for transmitting orders, weather reports etc., the speech's individual character is of little consequence. By reducing the bandwidth we have thus managed to remove the *redundancy* in the message. Only rarely will all the redundancy be removed since this will help to increase the comprehensibility if there are disturbances. Similar considerations can be made for other communication systems, for example for telegraphy and television. In the latter, transmitting the entire visible spectrum (3×10^{14} p/s) has never been suggested, but the image is analysed, a code signal is transmitted and the image is recovered in the receiver by synthesis; this is all in full analogy with the previously discussed vocoder telephony. The conclusion is that *communication systems must be constructed based on knowledge of the information source, the destination and what is to be transmitted*. If information is being transmitted to or from a person we need to recognise the reaction of different messages for the type of people that will use the system. Work is therefore currently underway in the USA to measure the comprehensibility and quality of messages to the ear and eye under different conditions.

Up to now we have used the expression *information* without further discussion of what it actually is. We have stated that we shall concentrate the attention not so much on the time function, as on *what* the time function represents, and this *what* is information. We have furthermore indicated that each message is selected from the class of messages that we wish to transmit; information therefore has something to do with *selection*. The difficulty in trying to give a closer definition of what we mean is not only due to the fact that we have included people in our communication system. Even in cases where text is transmitted from paper roll to paper roll in telegraphy (where people are not counted as either information sources or destinations), we have not yet been able to account for how much information can be transmitted per second. Neither have we been able to say how much information is lost if 1% of the characters are transmitted incorrectly. We therefore need a measure (i.e. a *unit*) of information, just like 1 volt is the unit for electric voltage. A unit of this type is defined⁵⁾ and has been given the designation 1 bit (or 1 Hartley).

With a unit of this type at our disposal we can logically attack the problem of transmitting information as effectively as possible. Let us consider a simple example from telegraphy. The message consists of letters of varying frequency of occurrence. In Norwegian and English the letter *E* appears more often than for example *Z*; in an efficient system we should therefore allow the signal for *E* to be shorter than the signal for *Z*, as in Morse code. In a telegraphy system for the Czech language, the messages will have other letter frequencies, and a different code is therefore desired. Similar considerations can be made for the other types of communication systems. What is generally involved is converting the message so that it *statistically adapts* to the channel. This process is known as *coding*, and will generally require a *time delay* in the sender, which consequently must be equipped with a memory of one type or another. In order to avoid misunderstandings, we draw your attention to the fact that this form of coding must not be confused with the previously discussed simplification of the message by removing the redundancy.

It can be shown⁶⁾ that a channel with bandwidth W p/s and signal-to-noise ratio K_S (reduced to sender-out) during the time T can transmit a quantity of information:

$$H = TW \log_2 (1 + K_S) \text{ bits}$$

With optimal coding, the channel's *information capacity* is thus

$$C = H/T = W \log_2 (1 + K_S) \text{ bits/sec}$$

C expresses the maximum number of bits per second that it is possible to transmit over such a channel. There is thus an *optimum* that can easily be calculated but which in practice can only be achieved with optimal coding. Generally speaking however, it is also possible to calculate how far from this optimum the individual systems are. Previously, a communication system's effectiveness for transmitting information could only be estimated by carrying out comprehensibility tests, which is a subjective measurement unless many different people are used to represent the information source and destination. Now on the other hand, we have obtained a tool for calculating (i.e. *objectively* calculate) a system's information efficiency.

From the expression above we see that for the same information capacity the bandwidth W can

be reduced, if K_S , i.e. the sender output, is increased sufficiently. Conversely, an increase in the bandwidth will allow a lower sender output. This concurs exactly with practical experience.

In the introduction we mentioned the three things that characterise a communication system: information, frequency band and signal-to-noise ratio. We have now arrived at a general *relationship* between them.

We have confined ourselves above to considering electrical communication systems. The points of view that have been expressed can however be used with regard to communication systems in a much wider context. The science that generally deals with the control and transmission of information in living beings and machines, has been given the name *Cybernetics* (Greek for steersman) by its founder, Norbert Wiener.

What is fundamentally new is that we have found telecommunications to be a statistical process and that a unit is defined for information. Whilst before we considered the different systems individually, we can now look at them collectively. This line of thought is very important; it opens up great possibilities and will undoubtedly gain greater practical importance as we define statistical characteristics for the different classes of messages we want to transmit.

Appendix

Correlation between "sinusoidal analysis" and "δ pulse analysis"

The system function $H(j\omega)$ for a network can be defined as the relationship between a signal's frequency spectrum at output $F_u(j\omega)$ and input $F_i(j\omega)$

$$H(j\omega) = \frac{F_u(j\omega)}{F_i(j\omega)}$$

We select a δ pulse as the signal; it can be regarded as the limit shape for the pulse

$$\frac{a}{\sqrt{\pi}} \exp(-a^2 t^2) \text{ when } a \rightarrow \infty, \text{ i.e. frequency}$$

spectrum-in, is in accordance with Fourier's integral formula

$$F_i(j\omega) = \int_{-\infty}^{+\infty} \lim_{a \rightarrow \infty} \left[\frac{a}{\sqrt{\pi}} \exp(-a^2 t^2 - j\omega t) \right] \cdot dt$$

or it is permitted to reverse the order of the limit and the integral,

5) "Information Theory", l.c.

6) "Information Theory", l.c.

$$F_i(j\omega) = \lim_{a \rightarrow \infty} \int_{-\infty}^{+\infty} \frac{a}{\sqrt{\pi}} \exp(-a^2 t^2 - j\omega t) \cdot dt = 1 \quad (2)$$

The network's δ pulse response is $h(t)$, i.e. frequency spectrum-out

$$F_u(j\omega) = \int_{-\infty}^{+\infty} h(t) \exp(-j\omega t) \cdot dt \quad (3)$$

Consequently, with the insertion of (2) and (3) in (1)

$$H(j\omega) = \int_{-\infty}^{+\infty} h(t) \exp(-j\omega t) \cdot dt$$

or according to Fourier's integral formula

$$h(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} H(j\omega) \exp(j\omega t) \cdot d\omega$$

There is therefore a unique relationship between $H(j\omega)$ and $h(t)$.

Literature

Shannon, C E. A Mathematical Theory of Communication. *Bell Syst. Techn. J.*, 27, 379–423, 623–656, 1948.

Shannon, C E. Communication in the Presence of Noise. *Proc. IRE*, 37, 10–21, 1949.

Wiener, N. *Cybernetics*. Wiley, 1948.

Wiener, N. *Extrapolation, Interpolation, and Smoothing of Stationary Time Series*. Wiley, 1949.

Halsey, R J, Swaffield, J. Analysis-Synthesis Telephony with Special Reference to the Vocoder. *Proc. Inst. Elec. Eng.*, III 95, 391–405, 1948.

Statistically Optimal Networks

Summary of lecture presented at a study session for radio technology and electroacoustics at Farris Bad, 16-18 June 1950
by Graduate Engineer Nic. Knudtson, Norwegian Defence Research Establishment, Bergen



Dr. Nic. Knudtson (80) obtained his Engineering degree from the Technical University of Norway, Trondheim in 1947 and his Doctor's degree from the Technical University in Delft, the Netherlands in 1957. 1948–1949 he was with the Research Laboratory of Electronics, Massachusetts Institute of Technology, working with information theory and experiments. 1950–1955 he was with the Norwegian Defence Research Establishment, Bergen, working on the development of microwave radio links; and from 1955 to 1967 he was Head of the Communications Division at Shape Technical Center in The Hague, Netherlands, where his efforts went into the planning of military telecommunications networks and systems in Western Europe. From 1968 to 1992 he was Director of Research at the Norwegian Telecommunications Administration, working on the planning of future telecommunications systems, networks and services. Dr. Knudtson has been member of government commissions and various committees, including the Norwegian Research Council, the National Council for Research Libraries, the International Telecommunications Union, EURESCOM, etc.

This is a translation into English of the paper "Statistisk optimale nettverk", which appeared in *Elektroteknisk Tidsskrift* 63 (30), 413–416, 1950. The translation was done by Berlitz GlobalNET and final quality control was done by Geir E. Øien.

A more detailed presentation will be issued as a separate publication. Only a brief outline of the problem will be presented here without any details about the mathematical solution.

The *individual* messages (signals) that are transmitted over a communication system cannot be known in advance if they are to contain any information for the addressee. The only thing that is known is certain *statistical characteristics* for the *class* of messages we want to transmit.¹⁾

Previously, electrical networks were constructed on the basis of messages that were a sine curve or a single δ pulse. The responses for these time series are determined forever, and will *not* transmit any information. These methods are not therefore based on realistic situations. However, all other time series can conceivably be built up of such sine curves or δ pulses. We have been aware of this for a long time, but until recently we have neglected to take account of the individual components' *statistical weight*, i.e. how often the individual frequency components or δ pulses appear in the messages. When we characterise an electrical network, we should combine the system function $H(j\omega)$ and the δ pulse response $h(t)$ with the statistical characteristics for the class of messages the network is to transmit. As we know, the following unique relationship exists between an electrical network's system function and δ pulse response

$$H(j\omega) = \int_{-\infty}^{+\infty} h(t)\varepsilon^{-j\omega t} dt$$

$$h(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} H(j\omega)\varepsilon^{j\omega t} d\omega$$

A filter separates a desired message from an unwanted message and noise. Previously, work was carried out in the *frequency domain*, and the filter's critical frequencies were set more or less randomly based on practical tests, where it was found which amplitude and phase response separated the unwanted messages from the desired messages *most effectively* without any *substantial* distortion. "Most effective" and "substantial" were however based on subjective assessments. The calculations, therefore, were not founded on any firm basis. Gradually the need for more complete solutions increased, and thus the television problems were created, which due to the nature of the message were different from those previously encountered in telephony and telegraphy. We were then led to proceed tentatively with finding usable solutions in this new field. Communication engineering was in many ways an *art*, since not only the principle of the constructions but their application depended a lot on the engineer's ingenuity and "good nature", whereas the quality could be discussed depending on personal taste. With the wealth of experience we have gradually gained in this way, very good and efficient networks have undoubtedly emerged. However, we have still not been able to judge *how* efficient these networks are, because the optimum has remained unknown. Moreover, we have been led to proceed tentatively each time a new type of problem has turned up. In brief, we have missed a general and well-founded theory for networks that will

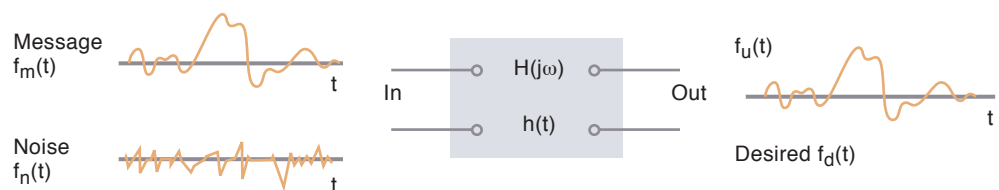


Figure 1 The network problem

¹⁾ See also "Statistical communication theory. A brief outline of the problem", *Teknisk Ukeblad*, 1950.

$f_n(t)$	$f_{\alpha}(t)$	network
$f_m(t) + f_n(t)$	$f_m(t)$	filter
$f_m(t)$	$f_m'(t)$	differentiator
$f_m(t)$	$f_m(t + \alpha)$	predictor
$f_m(t) + f_n(t)$	$f_m(t + \alpha)$	filter-predictor

transmit information of different kinds. The problem and methods we are now dealing with constitute a step in the development towards this goal.

We will now formulate the network problem in the *time domain*, see Figure 1. The messages are regarded as being continuous functions of time, where the redundancy is removed to the extent we wish. The information in a message of this type can easily be calculated.²⁾ Our problem is now to create a network in such a way that for certain classes of time functions-in $f_i(t)$ we get time functions-out $f_u(t)$ which deviate as little as possible from the *desired* time functions $f_d(t)$. The deviation is defined in the time domain as the mean squared error, i.e.

$$E = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^{+T} [f_u(t) - f_d(t)]^2 dt = \text{minimum} \quad (1)$$

Networks that are constructed based on this minimum condition are defined as *statistically optimal*. Some examples of these networks are given in the table below, where $f_m(t)$ represents a message and $f_n(t)$ represents noise or an unwanted message – both have known statistical characteristics.

It is worth noting that we can specify a time shift α in the time domain, and are thereby able to construct networks, known as predictors, which can predict the outcome for $f_m(t)$.

It is assumed in the following that

- 1) all time functions-in are *statistically stationary*, i.e. their statistical characteristics are invariant upon a translation in time of the associated time function. Commonly occurring messages and noise will most often fulfil this condition.
- 2) the network is *linear*. This condition is enforced due to the mathematical difficulties when solving non-linear problems. A solution of the linear problem is however better than none at all.

As we already know, the following relationship exists between time functions-out $f_u(t)$ and time functions-in $f_i(t)$ in a linear electrical network³⁾

$$f_u(t) = \int_{-\infty}^{+\infty} f_i(t - \tau)h(\tau)d\tau \quad (2)$$

where $h(t)$ is the δ pulse response.

Upon insertion of equation (2) in (1), the condition for obtaining a minimum is attained on the following form

$$E = \int_{-\infty}^{+\infty} h(\tau)d\tau \int_{-\infty}^{+\infty} h(\sigma)d(\sigma) \left[\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^{+T} f_i(t)f_i[t - (\sigma - \tau)]dt \right] \cdot 2 \int_{-\infty}^{+\infty} h(\tau)d\tau \left[\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^{+T} f_i(t - \tau)f_d(t)dt \right] + \left[\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^{+T} f_d(t)f_d(t)dt \right] = \text{minimum} \quad (3)$$

The expressions in the brackets are all of the form

$$\varphi_{12}(\chi) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^{+T} f_1(t)f_2(t \pm \chi)dt$$

This function is known as the *correlation function*; for $f_1(t) = f_2(t)$ we get *auto-correlation*, and for $f_1(t) \neq f_2(t)$ we get *cross-correlation*.

The auto-correlation function $\varphi_{11}(\chi)$ is a very important statistical parameter for the associated time function $f_1(t)$. Its most important property is expressed in Wiener-Khintchine's theorem: $\varphi_{11}(\chi)$ is the Fourier transformation of the *power density spectrum* $\phi_{11}(j\omega)$, which is defined as the mean output per frequency unit for $f_1(t)$. This is expressed mathematically as follows:

$$\Phi_{11}(j\omega) = \int_{-\infty}^{+\infty} \varphi_{11}(\chi) \cos \omega\chi d\chi \quad (3a)$$

$$\varphi_{11}(\chi) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \Phi_{11}(j\omega) \cos \omega\kappa d\omega \quad (3b)$$

Corresponding relations apply to the cross-correlation function. Equation (3a) forms the key equation for the *correlation method*: the power density spectrum can be determined via the correlation function, which is often easily calculated based on its definition equation by means of statistical methods.

³⁾ See for example "Theory of Servomechanisms", R.L. Series No. 25, p 35.

²⁾ "Information theory", lecture presented at a study session for radio technology and electro-acoustics at Farris Bad, 16–18 June 1950. ETT no. 30, 1950.

Upon insertion of the correlation functions, equation (3) takes the following form:

$$E = \int_{-\infty}^{+\infty} h(\tau) d\tau \int_{-\infty}^{+\infty} h(\sigma) d(\sigma) \varphi_{ii}(\tau - \sigma) - 2 \int_{-\infty}^{+\infty} h(\tau) d\tau \varphi_{id}(\tau) - \varphi_{dd}(0) = \text{minimum} \quad (4)$$

Here the correlation functions are known for the classes of time series we will deal with, and the network is selected so that its δ pulse response $h(\tau)$ makes the error E minimum. This is a problem of variation, which is solved by varying $h(\tau)$ and using Euler-Lagrange's condition for extremal values. Equation (4) will hereby be reduced to Wiener-Hopf's integral equation

$$\int_{-\infty}^{+\infty} h(\sigma) d\sigma \varphi_{ii}(\tau - \sigma) - \varphi_{id}(\tau) = 0 \quad \text{for } \tau > 0 \quad (5)$$

When solving this equation, consideration is given to the fact that the network must be *stable*. The condition for this is $h(\tau) = 0$ for $\tau < 0$, or expressed in the frequency domain, that the system function $H(\lambda)$ does not have poles inside the right half plane.

The system function for the statistically optimal network can then be shown to be

$$H(\lambda) = \frac{1}{2\pi \Phi_{ii}^{(V)}(\lambda)} \int_0^{\infty} \varepsilon^{-\lambda t} dt \int_{c-j\infty}^{c+j\infty} \frac{\Phi_{id}(w)}{\Phi_{ii}^{(H)}(w)} \varepsilon^{wt} dw \quad (6)$$

where $\lambda = \sigma + j\omega$ and $w = x + jy$ are complex-valued variables. The power spectrum for time series-in $\Phi_{ii}(\lambda)$ is split into factors $\Phi_{ii}^{(H)}(\lambda)$ and $\Phi_{ii}^{(V)}(\lambda)$, which have all poles and zeros in the right and left half respectively, i.e.

$$\Phi_{ii}(\lambda) = \Phi_{ii}^{(H)}(\lambda) \cdot \Phi_{ii}^{(V)}(\lambda)$$

When $H(\lambda)$ is thus determined, we will find a configuration of common network elements that has this system function. This is known as *synthesis* and there are several methods for solving this problem. It must be pointed out that the synthesis problem does not have a unique solution; there are several different configurations that have one and the same system function.

It is also possible to find an expression for the size of the error. Just like the system function, this is completely and exclusively determined by the correlation functions (or their Fourier transforms, which are the power spectra) for the classes of time functions that occur at input and which are desired.

As an example we will find the special system function for a *predictor*. Here is

$$f_i(t) = f_m(t) \text{ and } f_d(t) = f_m(t + \alpha)$$

where α is the prediction time.

Consequently

$$\Phi_{ii}(\lambda) = \Phi_{mm}(\lambda)$$

and it can further be shown that

$$\Phi_{id}(\lambda) = \Phi_{mm}(\lambda) \varepsilon^{\lambda \alpha}$$

so that, after equation (6)

$$H(\lambda) = \frac{1}{2\pi \Phi_{mm}^{(V)}(\lambda)} \int_0^{\infty} \varepsilon^{-\lambda t} dt \int_{c-j\infty}^{c+j\infty} \Phi_{mm}^{(V)}(w) \varepsilon^{w(t+\alpha)} dw \quad (7)$$

We will apply this equation to a simple example. Let $f_m(t)$ be a time series as shown in Figure 2, consisting of a series of identical pulses of the form $t\varepsilon^{-t}$. The single pulses are totally independent of each other and are random, there are k per unit time. The correlation function can be shown to be

$$\varphi_{mm}(\tau) = \frac{k}{4} \varepsilon^{-|\tau|} (1 + |\tau|)$$

and by applying equation (3a) the power spectrum is found as

$$\varphi_{mm}(j\omega) = \frac{k}{2\pi} \frac{1}{(1 + \omega^2)^2}$$

This is factorised as follows

$$\varphi_{mm}(j\omega) = \frac{k}{2\pi} \frac{1}{(1 + j\omega)^2} \frac{1}{(1 - j\omega)^2}$$

and this results in

$$\Phi_{mm}^{(V)}(w) = \sqrt{\frac{k}{2\pi}} \frac{1}{(1 + w)^2}$$

Figure 2 Signal

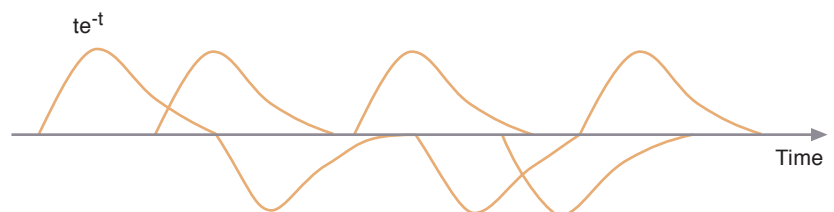


Figure 3 Predictor circuit for the signal in Figure 2

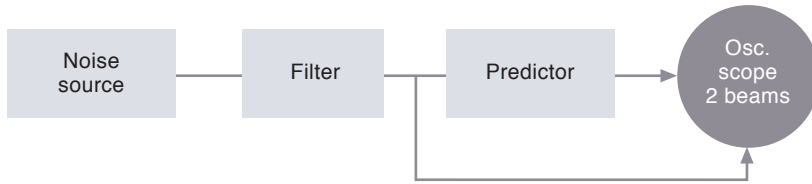
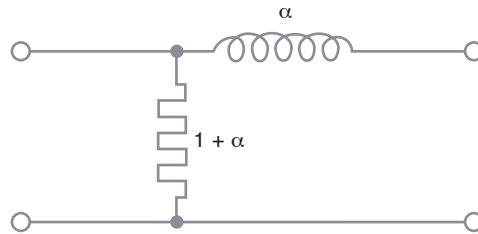


Figure 4 Demonstration of predictor for noise

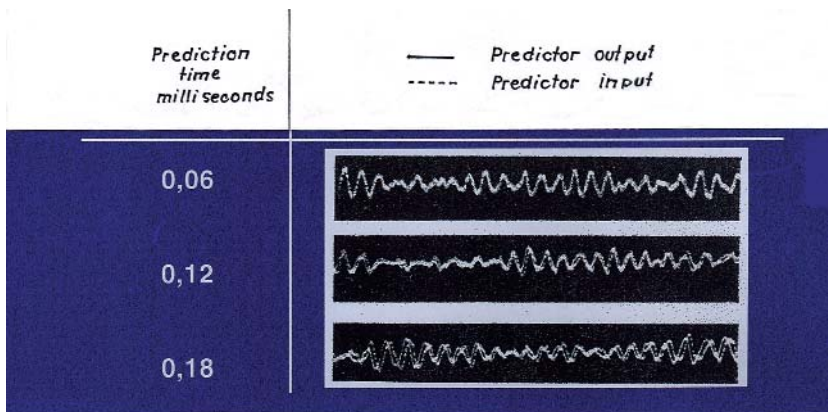


Figure 5 Example of prediction of filtered fluctuation noise

The system function for the optimal predictor for this time series is then

$$H(\lambda) = \frac{1}{2\pi} (1 + \lambda)^2 \int_0^{+\infty} \varepsilon^{-\lambda t} dt \int_{c-j\infty}^{c+j\infty} \frac{\varepsilon^{w(t+\alpha)}}{(1+w)^2} dw$$

or calculated, as the complex integration is performed using the residue theorem, as

$$H(\lambda) = \varepsilon^{-\alpha} (1 + \alpha + \alpha\lambda)$$

This is a network as shown in Figure 3. No guidelines have previously existed for the construction of predictors.

Lee & Stutt [3] have carried out a practical demonstration of predictors for filtered fluctua-

tion noise, see the block diagram in Figure 4. The noise source was a radio tube 6D4, and the filter a single resonance circuit of $f_o = 1080$ p/s and Q varying in stages from 10 to 90. The auto-correlation function for such filtered noise can be calculated; it will be a damped cosine function with frequency f_o and attenuation proportional to Q . The optimal predictor was then constructed based on Equation (7). Figure 5 shows an example of the prediction when $Q = 10$, the solid curve represents the time function at predictor-out and the dotted one at predictor-in. This results in the predicted function following nicely the actual function. Such predictors can be envisaged to have a bearing on the suppression of noise in communication systems.

Similar examples can be found for other types of statistical optimal networks: filters, differentiators, compensators etc, which are determined by Equation (6) upon due specification of the desired time series $f_d(t)$. This equation has therefore a very general validity.

Literature

- 1 Wiener, N. *Extrapolation, Interpolation and Smoothing of Stationary Time Series*. The Technology Press, John Wiley, 1949.
- 2 Lee, Y W, Wiesner, J B. Correlation Functions and Communication Applications. *Electronics*, 23, 86–92, 1950.
- 3 Lee, Y W, Stutt, C A. Statistical Prediction of Noise. *Proc. Nat. Elec. Conf.*, Chicago, 5, 342–365, 1949.

Excerpts from the Discussion After the Lecture

Garwick:

1. Derivation of the system function $H(j\omega)$ based on the delta function is not mathematically correct. The delta function is not a mathematical function; it does not satisfy Dirichlet's definition of a function, and cannot be integrated, either as a Riemann or a Lebesgues integral. When using delta functions, the order of the limit and integral sign are swapped. The limit must be carried out before the integral sign. It is possible that the result will be correct regardless, but this must be checked using other methods.
2. Where the error size E was discussed, there was absolutely no mention of the transmission channel's noise. Is the intention to imagine that the signal plus noise at input are already strengthened so much that the channel's set noise is negligible?

3. The equation:

$$E = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^{+T} [f_m(t) - f_d(t)]^2 dt$$

cannot always be used. $f(t)$ will in practice be almost equal to zero for $t > T_1$ and $t < T_2$. The integral can then be divided in three:

$$\frac{1}{2T} \int_{-T}^{+T_1} []^2 dt + \frac{1}{2T} \int_{T_1}^{T_2} []^2 dt + \frac{1}{2T} \int_{T_2}^T []^2 dt$$

When T goes to infinity, this expression goes to zero regardless of the network.

Knudtzon:

1. Communication engineers and mathematicians will probably never agree on the δ function. I refer to a very sound and worthy article by van der Pol.⁴⁾ I only used the δ function

when recapitulating *known* results from the network theory and could not take time to discuss this in more detail here on the blackboard, so as to avoid displacing the topic of my lectures.

2. *Linear* networks are assumed. The noise from the network is reduced therefore to the input.
3. *Statistically stationary time series* are expressedly assumed. The equation is mathematically correct.

C.B. & H.G.W.

⁴⁾ Philips Research Reports 1948, pp 174–190.

Multiple Bottom Lines?¹⁾ Telenor's Mobile Telephony Operations in Bangladesh

ARVIND SINGHAL, PEER J. SVENKERUD AND EINAR FLYDAL



Dr. Arvind Singhal (39) is Presidential Research Professor and Scholar in the School of Interpersonal Communication, College of Communication, Ohio University, where he teaches and conducts research in the areas of diffusion of innovations, mobilizing for change, design and implementation of strategic communication campaigns, and the entertainment-education communication strategy. Dr. Singhal is co-author of four books and has won several Top Paper Awards. He has won the Baker Award for Research at Ohio University twice, and numerous other teaching and research recognitions. Dr. Singhal has served as a consultant to the World Bank and UN programs, as well as private corporations. singhal@ohiou.edu



Peer J. Svenkerud (36) is Director of Stakeholder Relations, Telenor ASA. Svenkerud holds a PhD in organizational communication with specific focus on intercultural issues from Ohio University, and has conducted postgraduate studies at Harvard Business School. He is author of numerous peer-reviewed articles in international journals, and top academic papers at international conferences. His research interests center around information diffusion campaigns and the social impact of new communication technologies. Svenkerud was formerly director of Burson-Marsteller in Oslo, and Assistant Professor at the University of New Mexico. [peer-jacob.svenkerud@telenor.com](mailto:jacob.svenkerud@telenor.com)

The present article distills lessons learned about sound business and corporate social responsibility practices from Telenor's participation in mobile telephony operations in Bangladesh. Telenor's mobile telephony operations in Bangladesh provide valuable lessons about how corporations can strategically make forays in uncharted markets, and, while doing so, creatively seek and meet multiple co-existing bottom lines. Telenor is financially, intellectually, and structurally "richer" through its Bangladesh venture, gaining significant new insights and experiences in doing business in "distant" geographies, developing new culturally-based "benchmarking" standards, and experimenting with new business models that integrate telephone "ownership" with "access". Further, by partnering in Bangladesh with an internationally-known socially-driven development organization (i.e. the Grameen Bank), Telenor has strategically gained global visibility in both corporate and social sectors.

The social responsibility of business is to increase profits.

Milton Freidman, Nobel-prize-winning economist (quoted in Hood, 1996, p. 16).

Corporate social responsibility is not an occasional activity. It is not like visiting a mausoleum once a year, or hearing a church sermon every Sunday. It has to be completely integrated with the corporation's business function.

Muhammad Yunus, Managing Director of the Grameen Bank, in a personal interview (May 2, 2001).

The mobile phone is like a cow. It gives me "milk" several times a day. And all I need to do is to keep its battery charged. It does not need to be fed, cleaned, and milked. It has now connected our village with the world.

Parveen Begum, owner and sole dispenser of mobile telephony services in Village Chakalgram, Savar Thana, Bangladesh, in a personal interview (May 2, 2001).

I want my fellow Americans to know that the people of Bangladesh are a good investment. With loans to buy cell phones, entire villages are brought into the information age. I want people throughout the world to know this story.

U.S. President Bill Clinton in an address during his meeting with members of the Village Phone Project in Dhaka, Bangladesh in March, 2000.

The present article investigates²⁾ the mobile telephony operations in Bangladesh of Telenor, the leading Norwegian telecommunication company. Telenor has forged a business and strategic social change partnership with the Grameen Bank, one of the best known development organizations in the world. A historical background on Telenor's involvement in Bangladesh is provided. Telenor's business and social accomplishments in Bangladesh are presented, highlighting how a corporation can pursue multiple bottom lines.



This grocery shop is now the village information hub: The village phone is here during opening hours, and otherwise with the shopkeeper's wife at home



Einar Flydal (53) is cand.polit. from the University of Oslo, 1983 in pol. science, and a Master of Telecom Strategy from the University of Science and Technology (NTNU), Trondheim, 2002. Apart from 1985–1993, when engaged in IT in education for the Ministry of Education and with small IT companies, Flydal has worked with Telenor in a variety of fields since 1983. He has also worked as a radio freelancer on minorities and music; on statistical indicators at the Chair of Peace and Conflict Research, Univ. of Oslo, and on work organisation on oil platforms at the Work Research Institute, Oslo. Present professional interests: environment, CSR, innovation, and ICT.

einari.flydal@telenor.com

Telenor Goes to Bangladesh

How did Telenor get involved in Bangladesh?

To fully answer this question, a little background on the Grameen Bank operations in Bangladesh is useful. The Grameen (rural) Bank, founded in Bangladesh in 1983 by Professor Muhammad Yunus, is a system of lending small amounts of money to poor women so that they can earn a living through self-employment. No collateral is needed, as the poor do not have any. Instead, the women borrowers are organized in a group of five friends. Each group member must repay their loan on time, while ensuring that other group members do the same, or else their opportunity for a future loan is jeopardized. This delicate dynamic between “peer pressure” and “peer support” among Grameen borrowers is at the heart of its widespread success (Yunus, 1999). By December 2001, the Grameen Bank loaned money to about 2.4 million poor women borrowers, and had an enviable loan recovery rate of 95 percent. The idea of micro-lending, based on the Grameen Bank experience, has spread throughout the world, and has everywhere proven effective in gaining a high rate of repayment of the loans.

In the mid-1990s, the Grameen Bank began discussions with various mobile telephony operators around the world, including Telenor, to accomplish its vision of placing one mobile phone in each of the 68,000 villages of Bangladesh. At that time, there was one telephone in Bangladesh for every 400 people, representing one of the lowest telephone densities in the world³⁾. There was virtually no access to telephony services in rural areas, where 85 percent of Bangladesh’s 130 million people lived. Professor Yunus realized that while it was not possible for each rural household to own a telephone, it was possible through mobile telephone technology to provide access to each villager. To operationalize his vision, Professor Yunus established a non-profit organization called Grameen Telecom.

Telenor CEO Tormod Hermansen was intrigued by Professor Yunus’ idea of the village telephone, and believed that Bangladesh, given it only had 500,000 fixed line telephones in urban areas, presented a significant business opportunity for Telenor. Mobile telephony services could address the large unmet demand for telephony in Bangladesh, where the waiting period for a private fixed line connection was ten years. While there were significant “first-mover” advantages to be gained, the business risk was extremely high, given the unpredictable nature of Bangladesh’s political and regulatory environment. Telenor had previously never conducted business in a developing country in Asia, and Bangladesh seemed aeons away from

Norway. While Mr. Hermansen’s top advisors were “torn” about whether or not to foray into Bangladesh, Mr. Hermansen was enthusiastic, and provided patronage for the project to move forward⁴⁾.

So, in 1996, Telenor and Grameen Telecom formed a joint venture company called GrameenPhone Ltd (GP). Telenor provided 51 percent of the equity investment, Grameen Telecom provided 35 percent, Marubeni of Japan provided 9.5 percent, and Gonofone Development Corporation of USA provided the balance 4.5 percent. The Company, GrameenPhone, was awarded license to operate nation-wide GSM-900 cellular network on November 11, 1996. GP started its operation on March 26, 1997.

In launching its Bangladesh operations, GrameenPhone knew that its commercial viability depended on meeting the large unmet need for telephony services in urban areas. Since its inception in 1997, GrameenPhone’s subscription has doubled each year to reach 500,000 subscribers by December 2001. The company turned a profit three years later in 2000, with even brighter business prospects ahead: Demand for mobile telephony services in Bangladesh is estimated at about 5 to 6 million subscribers (out of a population of 130 million people). GrameenPhone’s growing mobile telephony network in the country, and its financial viability, helps the Grameen Telecom’s Village Phone Project to piggyback on it.

GrameenPhone sells air time in bulk to Grameen Telecom to re-sell it to members of Grameen Bank in villages. The eventual goal is for one Grameen borrower in each of the nation’s 68,000 villages to become the “telephone lady” for her village. Some 10,000 villages have been covered until February 2002. The village telephone lady operates a mobile pay phone business, with the cheapest cellular rate in the world: 9 cents per minute during peak hours and 6.7



More than 10,000 village phone ladies can show you with pride their new means of income and social advancement: the mobile phone

cents in the off-peak. Her “mobile” presence means that all village residents can receive and make telephone calls, obviating the need to install expensive large-scale telephone exchanges and digital switching systems.

Strategic Importance of GrameenPhone

The Telenor – GrameenPhone venture is of tremendous strategic importance to Telenor for at least two compelling reasons:

- 1 *GrameenPhone was a majority stake, start-up venture for Telenor*, as opposed to it purchasing a minority holding in an already established telecommunications business venture (as is the case with Telenor’s involvement in Thailand and Malaysia). So Telenor was involved in launching GrameenPhone from day one, thus experiencing the full gamut of pioneering experiences.
- 2 *The Bangladesh venture, metaphorically speaking, was as “distant” as could be from Telenor’s past business ventures.* Here was an established, affluent, Norwegian corporation, on the cutting-edge of telecommunications technology, adept at doing business in a stable political and regulatory environment, and in a country where the telephone density is the highest in the world, establishing a business venture in a far-away, fledgling, unserved and underserved Bangladeshi market, where the telephone density is about the lowest in the world, and the political and regulatory environment is relatively unstable. When one adds to this the social and cultural “distance” between Norwegians and Bangladeshis, one realizes this venture was bound to yield significant new learnings for Telenor.

Multiple Bottom Lines

Telenor’s involvement in the GrameenPhone mobile telephony project in Bangladesh has yielded an impressive list of business and social accomplishments:

Market Penetration and Return on Equity

- As noted previously, GrameenPhone’s subscription has more than doubled each year to reach 500,000 subscribers by December, 2001, which represents the biggest subscriber base and coverage of any mobile telephony operator in Bangladesh, and in the entire South Asia region. Mobile telephony users⁵⁾ (650,000) in Bangladesh in December 2001 outnumber the country’s fixed-line telephone subscribers (590,000).
- GrameenPhone earned a profit in both 2000 and in 2001 and is strategically positioned in 2002 for “explosive” growth. The company’s

market value now, estimated by its management at a modest \$600 to \$800 per subscriber, is \$300 to \$400 million (U.S.). These numbers suggest that GrameenPhone’s present value to Telenor is about six to seven times its majority (51 percent) equity investment of \$40 plus million.

Contribution to the Bangladesh Economy

- GrameenPhone, to date, has invested \$160 million in Bangladesh, making it the largest foreign private investor in the country.
- GrameenPhone has to date contributed \$75 million to the national treasury of Bangladesh in the form of new telephony tariffs, license fees, fees for leasing of the fiber-optic line, and other such receivables.
- By February 2002, GrameenPhone has directly created about 600 jobs (its employee strength) internally and 10,000 jobs externally in 10,000 Bangladeshi villages through the Village Phone Project of Grameen Telecom.

Service Provision in Unserved and Underserved Areas

- The 10,000 village-based mobile phones, leased or purchased by women members (also called “village telephone ladies”) of the Grameen Bank through a loan, serve 18 million rural inhabitants, who previously did not have access to telephony services. By the end of 2004, the number of village phones will likely increase to 40,000, serving an estimated 75 to 80 million rural inhabitants, about two-thirds of the entire Bangladesh population.
- The village phones, on average, generate 3–4 times more revenues for GrameenPhone than an individual use city/township subscription.
- The village telephone ladies, on average, make \$70 to \$80 per month of net profit from selling mobile telephony services in rural areas, which amounts to three times the per capita GNP of Bangladesh.
- In overall terms, the Village Phone Project (VPP) makes telephony services accessible and affordable to poor, rural Bangladeshis, spurs employment, increases the social status of the village telephone ladies, provides access to market information and to medical services, and represents a tool to communicate with family and friends within Bangladesh and outside. Studies indicate that the VPP has had a tremendous positive economic impact in rural areas, creating a substantial consumer surplus, and immeasurable quality-of-life benefits (Richardson et al., 2000; Bayes et al.,

In towns, the phone service shop may often be split between a front office for men, and a back entrance for women. The phone lady in front of her books show people that gender patterns can be changed



1999). For instance, the village phone now obviates the need for a rural farmer to make a trip to the city to find out the market price of produce, or to schedule a transport pick-up. The village phone accomplishes the task at about one-fourth the transportation costs and almost instantaneously (as compared to the hours of time it can take to make the trip), serving as a boon to the rural poor.

Gains in Intellectual and Structural Capital

What gains in intellectual and structural capital have accrued to Telenor through its involvement in GrameenPhone?

Pioneering Experience in Uncharted Markets

- *Reaching out to a relatively underserved and unserved telecommunications markets, while relatively risky for an organization like Telenor, is a viable business proposition especially if it wishes to expand operations geographically, and pioneer in gaining new business competencies. Being an early entrant in the field of mobile telephony in a relatively unserved and underserved telecommunications market bodes well for Telenor to maintain its dominant market leader position in Bangladesh, and to forge new opportunities elsewhere.*

Rethinking Benchmarks

- *Telenor has learned that conventional European benchmarking for estimating market potential – using measures of per capita GNP or Western patterns of telecommunications traffic – may be inappropriate or at least inadequate in the Asian context. In Bangladesh, for instance, people spend a much higher percent of their disposable income on*

telephony than in Western countries. Also, Asian cultures, by virtue of their “collectivistic” orientation and extended kinship structures, spur more frequent telephone talk between family members and friends, and for extended durations.

Teleaccess Business Models

- *In relatively underserved and unserved telecommunication markets, business potential should be evaluated not just on the basis of teledensity (or potential ownership) but also on teleaccessibility (or potential for providing access to those who cannot afford to own a subscription), as evidenced by the large reach of the Village Phone Project.*

Choice of Partner

- *In relatively “uncharted” markets or new geographies, it is of critical importance to choose a suitable local partner who adds long-term complementary value. Telenor’s major partner in Bangladesh is Grameen Telecom, a non-governmental organization floated by the internationally-acclaimed microlending institution, Grameen Bank, which has 2.4 million borrowers in 41,000 of the 68,000 Bangladeshi villages. In Bangladesh (and overseas), “Grameen” has tremendous brand equity by virtue of its widespread success in poverty alleviation, empowerment of rural women, and its well-known credo that “good development is good business” (the slogan of the the Village Phone Project).*

Many in Bangladesh feel that the “Grameen” brand is far more recognized in Bangladesh than even Coca Cola! So branding the new venture “GrameenPhone” brought instant credibility to Telenor’s business venture in Bangladesh. Also, Telenor’s partnering with Grameen Telecom made possible the Village Phone Project, whereby Grameen Bank borrowers who take loans to lease or purchase the mobile telephone sets now settle their monthly telephone bills through the bank workers. The already existing village-based loan disbursement and repayment infrastructure of the Grameen Bank allows for handling the logistics of the Village Phone Project at a very small, additional marginal cost. As noted previously, while the Village Phone Project represents some 10,000 subscribers (2 percent of GrameenPhone’s 500,000 subscribers), and a relatively small percent of its revenues (6 to 8 percent), it yields very high social impact in terms of reaching 18 million rural Bangladeshis who previously did not have access to telephony services, resulting in significant quality-of-life enhancements for them.



- In an “uncharted” market or “new geography”, a local partner can also play a significant role in familiarizing an organization like Telenor with the various political, regulatory, social, and cultural uncertainties, and helping to cope with them.

Two-Way Learning

- Telenor’s foray into Bangladesh has not been a one-way flow of capital, technology, and organizational structures from Norway to Bangladesh. Rather, technology-transfer, knowledge sharing, and capacity building have occurred in both directions, accruing significant benefits for Telenor. Telenor has helped create a corporate culture at GrameenPhone that is perceived by its Bangladeshi employees as being democratic, relatively non-hierarchical, merit-based, and gender-sensitive. Also, Telenor pioneered in Bangladesh the idea of integrating health, safety, and environmental issues in its business practices: For instance, it has immunized all its employees against the Hepatitis-B virus, and fields a doctor in its corporate office who provides medical consultation to employees and regularly conducts training programs on occupational and health safety. In turn, employees of GrameenPhone in Bangladesh have pioneered several operational innovations in Bangladesh that hold tremendous value for Telenor in its greenfield companies and other established markets. For instance, the Customer Relations Division of GrameenPhone, in-house, developed software that cuts down the customer phone activation procedure from 19 computer keystrokes to two keystrokes. This innovation has significantly enhanced employee productivity, obviating the need for hiring additional

personnel. One employee can now activate up to 2,000 telephone subscriptions a day, as compared to a paltry 150 previously. Telenor is presently sharing this GrameenPhone customer activation software, through a CD-ROM, in other geographies.

Human Resources for a Global Marketplace

- Over 50 Telenor officials have spent varying periods of time in Bangladesh, gaining invaluable experience in living and conducting business in a foreign nation’s political, regulatory, bureaucratic, social, and cultural environment. Such experiences, laced with all kinds of uncertainty, adjustment, acculturation, and new learnings, contribute significantly to a corporation’s human resources in a global playing field.

Lessons for Business and Corporate Social Responsibility

In addition to an impressive list of commercial and social accomplishments in Bangladesh, tremendous public relations and promotional benefits accrue to Telenor by cooperating with the Grameen family of companies, which represent an icon of development organizing to the outside world. When the world’s leading agenda-setter, the U.S. President (at that time, Bill Clinton) visits with the village telephone ladies in Bangladesh and hails the integrated business and social aspects of their venture (as expressed in the statement listed at the top of this case study), mass media, policy-makers, corporations, and the public all over the world take notice.

Women’s network group leaving the Grameen Bank weekly micro-repayment meeting. Grameen Bank village facilities: building in corrugated iron sheets – itself a symbol of progress

Specifically, Telenor's foray into Bangladesh highlights the following lessons for its business and corporate social responsibility functions:

#1 Sound business means subscribing to multiple, co-existing, and mutually-reinforcing (win-win) bottom-lines, which also implies acting as a socially responsible corporation. In Bangladesh, Telenor's multiple bottom-lines included:

- *Meeting commercial interests* in terms of revenues, profits, and growth.
- *Meeting social cause-related interests* in terms of serving unserved and underserved markets nationally, and also serving poor, rural, illiterate inhabitants who are traditionally excluded from traditional markets (thus overcoming the digital divide).
- *Gaining substantial amount of experience in overseas operations* by doing business in a "distant" geography and an unfamiliar market, which helps build intellectual and structural capital for future ventures.
- *Gaining in "image" and "prestige"* by partnering in a unique commercial and social experiment with an internationally acclaimed branded local partner, the Grameen Bank. The "value" of endorsements from such world luminaries as U.S. President Bill Clinton, or the "value" of the GSM Community Award bestowed on the Village Phone Project during the GSM World Congress in Cannes, France, in 2000, are hard to gauge in pure economic terms, pointing to the value of recognizing multiple bottom-lines.

#2 Corporate social responsibility does not mean merely "showcasing" one initiative (such as the Bangladesh case), but rather

integrating CSR as a competitive asset in all business ventures. Hence, true corporate social responsibility means integrating all business functions with a social imperative, and measuring the effectiveness of the CSR function not just by what Telenor has achieved to date in Bangladesh, but what more can it achieve in the long-term.

At the present time, Telenor's operations in Bangladesh have centered around only one of its core competencies, i.e. mobile telephony. However, every aspect of Telenor's business (Internet services, communication satellites for narrow and broadband services, interactive Web-based services, cable television, telemedicine, fixed and mobile telephony services, and others) holds a strategic business potential in Bangladesh, and in other developing country markets. How can this business potential be tapped and leveraged in new geographies?

Telenor's well-established partnership with the Grameen Bank – which has launched several new information technology companies such as Grameen Telecom, Grameen Communications, Grameen Software, Grameen Cybernet, Grameen Shakti (power), and others – positions Telenor, like no other corporation in the world to experiment with new initiatives in E-health, E-education, E-commerce, E-banking, and other services that may have a long-term business as well as a social function. For instance, GrameenPhone's has 1,800 kilometers of available optical fiber (leased from Bangladesh Railways), which to date has been barely utilized. Can Telenor leverage its relationship with the Grameen family of companies to develop new business ventures with a social imperative?

To profit further on the lessons learned, should Telenor seriously consider looking at Bangladesh as the prime location for establishing an independent R&D and/or a Business Develop-

GrameenPhone's main income source is the cities. Here publicity boards in one of Dhaka's most fashionable hotels, shown us by GrameenPhone's information officer, Yamin Bakht





A new initiative: One of several Grameen IT education centers, providing the resources for self reliant software development and distance services like secretarial, programming, punching etc.

ment Center, mainly hiring talented Bangladeshi personnel to experiment with new initiatives in E-health, E-commerce, E-education, E-banking, and other need-based applications to establish ventures in one of the most “unserved” and “underserved” world markets? With its existing physical presence in Bangladesh through GrameenPhone, an already established relationship with a branded local partner in the Grameen Bank, and a tremendous base of already-gained intellectual and structural capital, can Telenor be at the forefront of developing new products and services which can be economically viable and also address the needs of unserved and underserved markets of Asia, Africa, and Latin America? Could Telenor, for instance, in association with the Grameen family of companies, experiment with the synergies that arise from the presence of credit (provided by Grameen Bank), connectivity (provided by GrameenPhone), and energy (provided by Grameen Shakti through solar panels) in unserved and underserved markets? With the microcredit movement growing by leaps and bounds around the world, and with the Grameen family of companies leading this march (Grameen replication efforts are now underway in over 75 countries), could Telenor position itself for new market opportunities as no other corporation?

In summary, Telenor’s mobile telephony operations in Bangladesh in co-operation with the Grameen system suggest that a corporation can strategically pursue multiple bottom lines. It does not imply that such a success story is always a clear and neat result of plans and detailed oversight. On the contrary, new options and innovative entrepreneurs do also imply trials and errors and new and uncommon problems to solve. None the less, Telenor’s operations in Bangladesh, especially the strategic integration

of business and corporate social responsibility functions, have resulted in an exemplary first mile in a long marathon race. Will Telenor consider running the full race?

References

- Bayes, A, von Braun, J, Akhter, R. 1999. *Village Pay Phones and Poverty Reduction*. Bonn, Germany, Center for Development Research.
- Hood, J. 1996. *The Heroic Enterprise*. New York, Free Press.
- Richardson, D, Ramirez, R, Haq, M. 2000. *Grameen Telecom’s Village Phone Programme in Rural Bangladesh : A Multimedia Case Study*. Ottawa, Canada, Canadian International Development Agency.
- Yunus, M. 1999. *Banker to the Poor*. New York, Public Affairs.

Endnotes

- ¹⁾ We thank the following individuals who helped us in implementing the present project: Tormod Hermansen, CEO of Telenor; Beth Tunland, Senior Vice President, Telenor’s Corporate Social Responsibility function; Marit Reutz, Director, Telenor Corporate University; Sigve Brekke, Managing Director, Telenor Asia; Ola Ree, Managing Director of GrameenPhone; Professor Muhammad Yunus, Managing Director of the Grameen Bank; Mr. Khalid Shams, Deputy Managing Director of the Grameen Bank; Syed Yamin Bakht, Additional General Manager of Information, GrameenPhone, and various others.
- ²⁾ Our data-collection procedures consisted of (1) extensive archival research, including reading of various evaluation reports on the GrameenPhone project (for instance, the Richardson et al., 2000;

and the Bayes et al., 1999 reports); books written on the Grameen Bank, including Professor Muhammad Yunus' (1999) book, *Banker to the Poor*, and others; (2) in-depth interviews at Telenor AS, Norway with key individuals involved in the planning and implementation of the Telenor-GrameenPhone project in Bangladesh, and with their Bangladeshi counterparts, including Professor Muhammad Yunus, Managing Director, and Mr. Khalid Shams, Deputy Managing Director of Grameen Bank; (3) a two-week field visit to Bangladesh for observation of, and in-depth interviews with key principals at GrameenPhone and the Village Pay Phone projects of Grameen Telecom. Our above field-based activities in Bangladesh yielded about 30 in-depth taped interviews, several volumes of field-notes, and over 250 photographs.

3) By December 2001, there is one telephone in Bangladesh for every 200 people, largely as a result of Telenor-GrameenPhone's mobile telephony operations.

4) Several middle and senior managers at Telenor continue to be worried about the sustainability of the GrameenPhone initiative. For the GrameenPhone business to continue growing, large capital investments are continually needed. In 2001, Telenor's other equity partners in GrameenPhone (Grameen Telecom, Marubeni, and Gonofone) were unable to raise their share of the new investments, which put Telenor in the awkward position of somehow raising (through credit) the needed funds at the last minute.

5) Some 150,000 mobile telephony users are served by other competitors of GrameenPhone.

Introduction

PER HJALMAR LEHNE



Per Hjalmar Lehne (44) obtained his MSc from the Norwegian University of Science and Technology in 1988. He has since been with Telenor R&D working with different aspects of terrestrial mobile communications. 1988 – 1991 he was involved in standardisation of the ERMES paging system in ETSI as well as in studies and measurements on EMC. His work since 1993 has been in the area of radio propagation and access technology, especially on smart antennas for GSM and UMTS. He has participated in the RACE 2 Mobile Broadband Project (MBS), COST 231, and COST 259 and is from 2001 vice-chairman of COST 273. His current interests are in 4th generation mobile systems and the use of MIMO technology in terrestrial mobile networks, where he participates in the IST project FLOWS.

per-hjalmar.lehne@telenor.com

In this issue of the Status section of *Telektronikk*, we focus on one of the most important but often underrated aspects of modern telecommunications, namely security. The section contains only one paper, however it is a very comprehensive description of UMTS Network Domain Security written by Geir M. Kjøien from Telenor R&D, who is a delegate in the 3GPP SA3 dealing with security.

In his paper, which is a follow-up from a previous paper in 2000, he addresses the further developments of the security specifications in UMTS Releases 4 and 5. Work on security for the control plane of the UMTS core network started with Release 4. Here only security of the Mobile Application Part (MAP) of SS7 was specified. This is referred to as MAPsec. Security for IP-based control plane protocols was scheduled for Release 5.

The paper thus contains a fairly detailed description of MAPsec, as it is specified by SA3 and given by one of the key people of this work. Additionally, the Network Domain Security for IP (NDS/IP) is explained and the limitations of the IP security protocol (IPsec) are addressed. He concludes with what lies ahead for UMTS Network Domain Security, namely the introduction of a Public Key Infrastructure (PKI) to support the use of digital certificates. Secure authentication methods are probably one of the most important mechanisms necessary to facilitate a wide use of e-commerce in general and m-commerce particularly.

UMTS Network Domain Security

GEIR M. KØIEN



Geir Køien (36) has been an employee of Telenor R&D since 1998. He has been working with various mobile systems since 1991 and is interested in security and signalling aspects of mobile systems. He is the Telenor delegate to 3GPP SA3 (Security) where he has served/ serves as rapporteur for the Network Domain Security specifications 3GPP TS 33.200 and 3GPP TS 33.210. He is also pursuing a PhD at Agder University College.

geir-myrndahl.koien@telenor.com

1 Introduction

This article is a follow-up to the article on *Overview of UMTS security for Release 99* in *Teletronikk 1.2000* [1]. During the last two years the UMTS security architecture has evolved to include security for the control plane of the core network as well as to cover security for the new IP Multimedia Subsystem (IMS) architecture.

This article will focus on describing the services and features of the core network control plane security extensions. These are collectively called Network Domain Security and comprise two technical specifications.

A brief description will also be given on the way forward for the UMTS security specification process.

1.1 Security Features for UMTS Release 4

The work to provide security for the control plane of the UMTS core network started for real with UMTS Release 4¹⁾. This work took place under the *work item* name of Network Domain Security (NDS) and the goal was to secure all important control plane protocols in the core network. This included both protocols based on the telephone signalling systems (SS7) protocol stack and protocols based on the IP protocol stack. It was realized that the SS7-based and IP-based protocols were sufficiently different to warrant separate security solutions. The work to protect IP-based protocols was scheduled for UMTS Release 5.

During the process for Release 1999 it was realized that to protect SS7-based protocols would inevitably mean to protect them at the application layer. The drawback of implementing protection at the application layer is that the target protocol itself will have to be updated in a pervasive and non-trivial way. This is an expensive and time consuming process that would have to be repeated for every target protocol. After careful deliberations it was found that one could not afford to protect more than a selected set of protocols and in the end it was decided that the only SS7-based protocol that would be protected was the Mobile Application Part (MAP) protocol [2].

The reason that one chose to protect MAP is that MAP is a crucial core network protocol that provides mobility management services and distributes the Authentication Vector (AV) security data from the HLR/AuC to the VLR/SGSN.

The technical specification 3GPP TS 33.200 *Network Domain Security; MAP application layer security* [3a] was completed for Release 4 by June 2001. The specification, which is often just called MAPsec, contained procedures for secure transport of MAP messages between MAP network elements, but lacked mechanisms for key negotiations and distribution. The key negotiation and distribution procedures are scheduled to be included in the Release 5 version of TS 33.200 [3b].

1.2 Network Domain Security Features for UMTS Release 5

The main Network Domain Security goals for UMTS Release 5 are:

- a) to provide Network Domain Security protection for IP-based control plane protocols;
- b) to complete the MAP security architecture to include key negotiation and distribution procedures.

The Network Domain Security for IP-based control plane protocols are based on the IETF IPsec protocols. The work was completed when the technical specification 3GPP TS 33.210 *Network Domain Security; IP network layer security* [4] was approved March 2002.

2 Network Domain Security; MAP Application Layer Security

In this section an attempt is made to explain the technical realization of the MAP security protection.

2.1 MAPsec Security Services

The security services provided by MAPsec are:

- Cryptographic data integrity of the MAP messages;
- Data origin authentication for the MAP messages;

¹⁾ The naming conventions for the releases changed and Release 4 is the subsequent release to Release 1999.

- Replay protection for the MAP messages;
- Confidentiality (encryption) for the MAP messages.

2.2 The MAP Security Architecture

The main MAP security architecture consists of the following elements and interfaces:

- **Key Administration Centre (KAC)**

- **Release 5**

- This new network element is responsible for key negotiation and distribution between network operators. The KAC is part of TS 33.200 Release 5 [3b].

- **MAP Network Elements (MAP-NE)**

- The MAP network elements must be updated to support the Zf-interface to participate in secure MAP communication. MAP-NEs conforming to MAPsec Release 5 specifications must also support the Ze-interface.

- **Zd-interface (KAC – KAC) – Release 5**

- The Zd-interface is an IP-based interface that is part of MAPsec Release 5. It is used to negotiate MAPsec Security Associations (SAs²) between MAP security domains. The only traffic over the Zd-interface is the Internet Key Exchange (IKE) negotiations of MAPsec security associations. The semantics of the MAPsec SAs are defined in *The MAP Security Domain of Interpretation for ISAKMP* informational RFC. At present only a draft version of the MAPsec DoI is available (draft-arkko-map-doi-04.txt [5]). The security services specified by the security association is encoded in the protection profile information element.

- **Ze-interface (KAC – MAP-NE) – Release 5**

- The Ze-interface is an IP-based interface that is part of MAPsec Release 5. This interface provides distribution of security association data from a KAC to a MAP-NE within one operator domain.

- **Zf-interface (MAP-NE – MAP-NE)**

- **Release 4**

- The Zf-interface is a MAP interface that is part of MAPsec Release 4. The MAP-NEs may be from the same security domain or from different security domains (as shown in Figure 1). The MAP-NEs use MAPsec security associations received from a KAC to protect the MAP operations. The MAP operations within the MAP dialogue are protected selectively according to the chosen MAPsec protection profile.

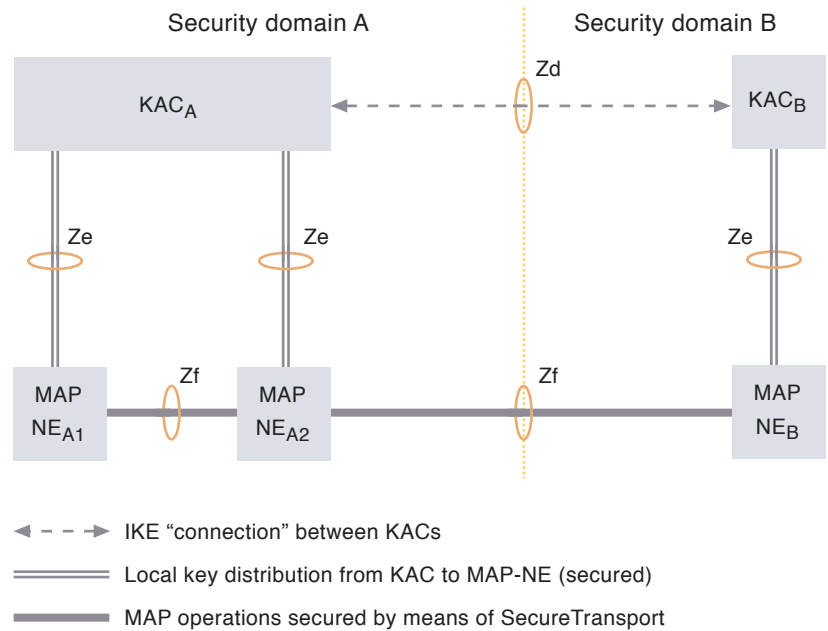


Figure 1 Overview of the Zd, Ze and Zf-interfaces (from TS 33.200 [3a])

For the purpose of the actual protection, the *SecureTransport* meta-operation component is used. The original MAP component is encapsulated in *SecureTransport*.

From an architectural point of view, one should note that the use of IKE for key negotiation introduces an anomaly. To use IKE for the Zd-interface may seem to be an obvious choice in that IKE is a protocol that is specifically designed to carry out key negotiations. However, IPsec/IKE make the basic assumption that IKE negotiates security associations on behalf of the network element on which it resides. For MAPsec this will *not* be the case since the KAC shall not use the MAPsec SA itself. MAPsec security associations are furthermore valid on a network-to-network basis and not individually between the communicating parties.

This means that one pair of security associations will be used for all MAPsec communication between two domains. In Figure 1 MAP-NE_{A1} and MAP-NE_{A2} will use the same security association pair when communicating with MAP-NE_{B1}. This also extends to the case were two MAP-NEs within the same security domain needs to engage in MAPsec secured dialogues, but here the security association (SA) pair must be a special initiator-SA_{self} and responder-SA_{self} pair.

² A MAPsec Security Association (SA) is a unidirectional logical control channel for a secured connection. The SA specifies the security services, the algorithms and the lifetime of the secured connection amongst others. Due to different requirements, a MAPsec SA is different from an IPsec SA.

Protection level	Protection mode for invoke component	Protection mode for result component	Protection mode for error component
1	1	0	0
2	1	1	0
3	1	2	0
4	2	1	0
5	2	2	0
6	2	0	0

Table 1 MAPsec protection levels (from TS 33.200 [3a])

2.3 MAPsec Protection Modes

MAPsec provides for three different protection modes. These protection modes regulate the protection level of the original MAP component:

- Protection Mode 0: No Protection**
 Protection Mode 0 uses the encapsulation provided by MAPsec, but offers no cryptographic protection.
- Protection Mode 1: Integrity, Authenticity**
 In Protection Mode 1 the protected payload is a concatenation of the cleartext and the Message Authentication Code (MAC) generated by the integrity function f7.
- Protection Mode 2: Confidentiality, Integrity, and Authenticity**
 Integrity in protection mode 2 is achieved by the same means as for protection mode 1. Confidentiality is achieved by encrypting the cleartext using the encryption function f6.

Note that in protection mode 0 no protection is offered and that the “protected” payload is identical to the payload of the original MAP message. However, since a protection mode 0 component is encapsulated by means of *Secure-Transport* it is not identical to the original component when it comes to the processing steps taken by MAP/MAPsec.

2.4 MAPsec Protection Profiles

The notion of a MAPsec protection profile was invented to simplify negotiation of security associations between roaming partners. The idea was to make agreements on the extent of MAPsec protection very simple, both qualitatively and quantitatively. This would make it much easier for roaming partners to create mutually compatible MAPsec security requirements. Unfortunately, the actual construction of the protection profiles has become somewhat counterintuitive, slightly inflexible and may appear a bit contrived.

Technically, MAPsec protection is specified per MAP operation component. The MAPsec protection profiles are organized by means of protection groups, which are loosely organized around the various transactions (dialogues) that MAP executes. Each protection group defines a set of MAP operations and their protection modes at the operation component level. The concept of “protection level” is introduced to administrate the protection mode on operation component level. A protection level of an operation determines the protection modes used for the operation’s components according to Table 1.

It shall be noted that not all MAP operations/components are included in the protection groups. The operations/components that are not present in any protection group cannot be protected by means of MAPsec.

The protection profiles are composed of non-overlapping protection groups and are predefined in TS 33.200. The current set of protection groups and protection profiles may be extended in later releases.

2.5 MAPsec Security Policies

MAPsec security policies are structured around the protection profile concept. The extent of security protection is thereby regulated both in terms of which operations to protect and the protection level by means of choosing the appropriate protection profile. The protection profile concept offers very simple policy management at the expense of flexibility.

Network operators must agree on which protection profile to use in bilateral agreements. These agreements should become part of the standard roaming agreements between operators.

3 Network Domain Security; IP network Layer Security (NDS/IP)

The most important IP-based protocol to protect in the UMTS core network control plane is the GTP-C protocol [6]. NDS/IP is defined on the network layer and it is therefore easy to adapt NDS/IP to protect GTP-C. In fact, no changes are required to the target protocol.

Technical specification *3GPP TS33.210 Network Domain Security; IP network layer security* [4] defines how the IP-based core network control plane protocols can be protected. The IETF already has a stable and well-defined architecture for IP security (IPsec, RFC-2401 [7]) at the network layer. It was therefore an obvious choice for SA3 to base NDS/IP on the security services afforded by IPsec. For the purpose of the closed domain of NDS/IP, many of the options and services of IPsec were redun-

dant. The use of IPsec in NDS/IP is therefore profiled to remove unnecessary functionality. The driving force behind the profiling was to reduce complexity in order to facilitate stable and effortless interoperability.

3.1 NDS/IP Security Services

IPsec offers a set of security services by means of the two security protocols Authentication Header (AH) (RFC-2402, [8]) and Encapsulating Security Payload (ESP) (RFC-2406, [9]). For NDS/IP the IPsec security protocol shall always be ESP and it shall always be used in *tunnel mode*. Tunnel mode is an IPsec mode that provides protection for the whole of the original IP packet and is typically used between security gateways. The other mode, *transport mode*, is targeted specifically towards end-to-end communication between users, and where the primary goal is to protect the payload portion of the original packet – i.e. providing protection for application data.

The security services provided by NDS/IP are:

- a) Connectionless data integrity
- b) Replay protection
- c) Data origin authentication
- d) Data confidentiality for the whole original IP packet (optional)
- e) Limited protection against traffic flow analysis when confidentiality is applied

List-1 Security services provided by NDS/IP

When using NDS/IP, the minimum protection level provided shall be integrity protection/message authentication with replay protection. This amounts to the services a), b) and c) in List-1. The ESP authentication mechanisms provide these security services.

Confidentiality protection (encryption) is an option when using NDS/IP, and for NDS/IP it shall always be used in conjunction with integrity protection as recommended in the ESP RFC [9]. The ESP confidentiality mechanisms are used to provide services d) and e).

NDS/IP also specifies the use of the Internet Key Exchange (IKE) (RFC-2409, [10]) protocol for automatic key negotiation and distribution.

3.2 IPsec Limitations

IPsec operates at the IP layer and provides its security services to the transport layer protocols. The transport layer protocols have traditionally been the User Datagram Protocol (UDP) [11] and the Transmission Control Protocol (TCP) [12]. Since IPsec is located at the network layer, its processing rules are exclusively based on information in the IP header. This means that IPsec cannot identify logical connection at

higher layers and cannot differentiate its services beyond the information found in the IP headers. For the purpose of the UMTS core network control plane protocols, this will not normally matter. The exception to this is that IPsec cannot be used to discriminate on the contents in IP tunnels. This means that NDS/IP, located at the UMTS control plane, cannot selectively protect the contents of the GTP-U protocol [6] since it cannot inspect the contents of the tunnel.

As mentioned above, IPsec would normally be used to provide security services to the transport protocols TCP and UDP. This is all fine, except that the new IP-multimedia services in UMTS will also use the new Stream Control Transmission Protocol (SCTP) transport layer protocol [13]. SCTP is amongst other things capable of supporting multiple IP addresses at each endpoint (multi-homing), and this is currently a problem with IPsec. For the time being one must therefore accept that IPsec, and therefore NDS/IP, cannot be guaranteed to support SCTP. Work is ongoing in the IETF to solve this problem, and NDS/IP will be updated to support SCTP when a solution is finalized for IPsec.

3.3 The NDS/IP Security Architecture

The NDS/IP key management and distribution architecture is based on the IPsec IKE protocol. This protocol provides automated Security Association (SA) negotiation and distribution for the IPsec protocols. During SA negotiation, IKE goes through two phases. Phase-1 is the set-up of the IKE “control channel” and phase-2 is where the actual IPsec SA negotiation takes place.

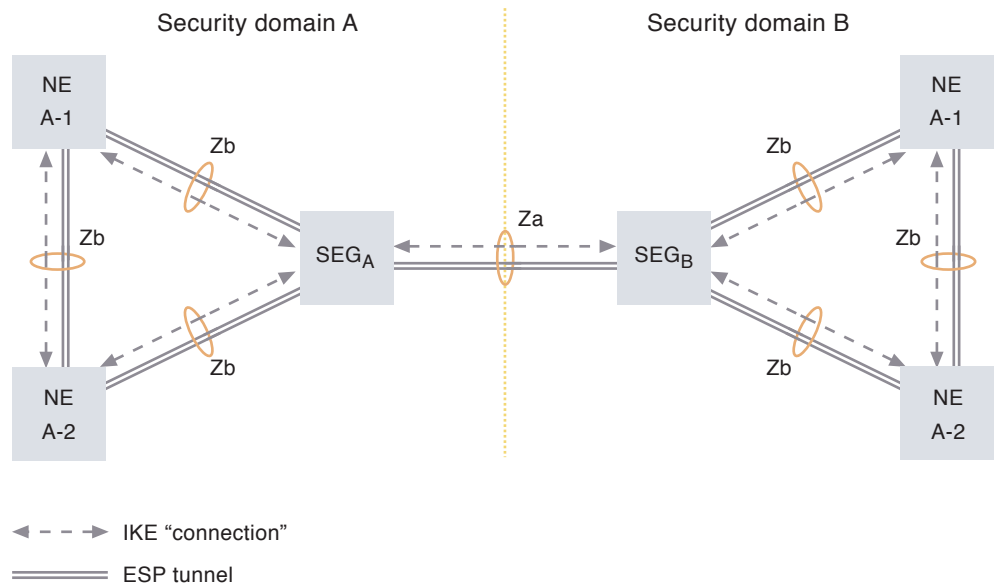
The basic idea of the NDS/IP architecture is to provide hop-by-hop security. This is in accordance with the *chained-tunnels* or *hub-and-spoke* models of operation. The use of hop-by-hop security also makes it easy to operate separate security policies internally and towards other external security domains.

In NDS/IP only the Security Gateways (SEGs) shall engage in direct communication with entities in other security domains for NDS/IP traffic. The SEGs will then establish and maintain IPsec secured ESP tunnels between security domains.

All NDS/IP traffic from an NE in one security domain towards an NE in a different security domain will be routed via an SEG and will be afforded *chained-tunnels* security protection towards the final destination.

Operators need only establish one ESP tunnel between two communicating security domains. This would make for coarse-grained security granularity. Alternatively, the operators may set

Figure 2 NDS architecture for IP-based protocols (from TS 33.210 [4])



up separate tunnels for each of the protocols and services that are protected³⁾.

The following interfaces are defined for protection of native IP-based protocols:

- **Za-interface (SEG-SEG)**
The Za-interface covers all NDS/IP traffic between security domains. The SEGs use IKE to negotiate, establish and maintain a secure ESP tunnel between them.
- **Zb-interface (NE-SEG / NE-NE)**
The Zb-interface is located between SEGs and NEs and between NEs within the same security domain. The Zb-interface is optional. Normally ESP shall be used with both encryption and authentication/ integrity, but an authentication/integrity only mode is allowed. The ESP tunnel shall be used for all control plane traffic that needs security protection. Whether the tunnel is established when needed or *a priori* is for the security domain operator to decide. The tunnel is subsequently used for exchange of NDS/IP traffic between the NEs.

The security policy established over the Za-interface is subject to roaming agreements. This differs from the security policy enforced over the Zb-interface, which is unilaterally decided by the security domain operator.

There is no NE-NE interface for NEs belonging to separate security domains. This is because it is important to have a clear separation between the security domains. This is particularly rele-

vant when different security policies are employed within the security domain and towards external destinations.

The restriction not to allow secure inter-domain NE-NE communication does not preclude a single physical entity to contain both NE and SEG functionality.

3.4 NDS/IP Encryption Algorithms

IPsec offers a wide set of confidentiality transforms. The mandatory transforms that compliant IPsec implementations must support are the ESP_NULL and the ESP_DES transforms. However, the Data Encryption Standard (DES) transform is no longer considered sufficiently strong in terms of cryptographic strength. For NDS/IP, neither of the mandatory transforms is allowed.

The new Advanced Encryption Standard (AES) [14] developed by NIST is expected to be available for IPsec shortly. For the purpose of NDS/IP, the ESP_AES transform(s) will be made mandatory when they are approved by the IETF. In the meantime the 3DES transform is to be used.

3.5 NDS/IP Integrity Algorithms

The integrity transforms that compliant IPsec implementation is required to support are the ESP_NULL, the ESP_HMAC_MD5 and the ESP_HMAC_SHA-1 transforms. Since NDS/IP traffic always requires the anti-replay service, the ESP_NULL transform is not allowed in NDS/IP.

³⁾ IPsec will normally discriminate traffic based on the quintuple (source-IP-address, destination-IP-address, source-port-number, destination-port-number and transport-protocol-identity).

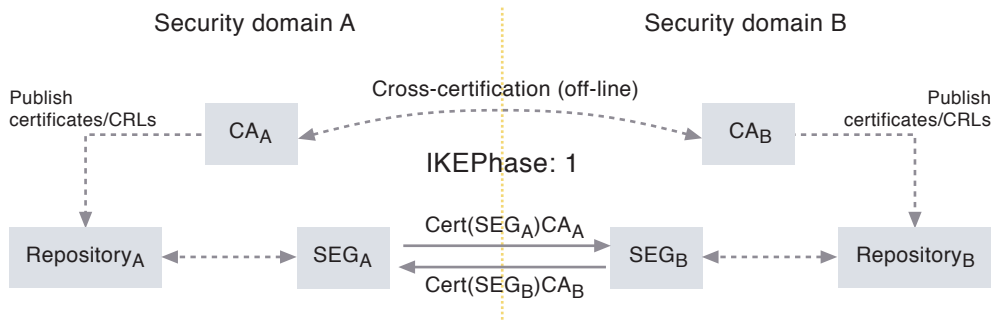


Figure 3 IPsec used with chain-tunnels, i.e. hierarchy still kept, one Certificate Authority (CA) per security domain which must be cross-certified with roaming partners (from S3-010622 [15])

The new Advanced Encryption Standard (AES) [14] developed by NIST is expected to be available for IPsec shortly. For the purpose of NDS/IP, the ESP AES MAC transform(s) will be made mandatory when they are approved by the IETF. NDS/IP implementations shall also support the ESP_HMAC_SHA-1 transform.

4 The Road Ahead for UMTS Network Domain Security

4.1 Scalability

For the IP-based services, a need is foreseen for an authentication framework for network elements. The basis for this assumption is that authentication for the IPsec IKE phase-1 “control channel” is currently provided by means of pre-shared secrets. Normally, the use of pre-shared secrets scales poorly. For NDS/IP the problem is manageable due to the inherent sectioning and hierarchy introduced by the security domains and the fact that protection is by means of chained tunnels.

Nevertheless, in an environment where the number of IP capable network elements is expected to rise dramatically, the need for a more scalable architecture is likely to be needed within the next few years. The rise in the number of IP addressable network elements is likely to coincide with migration to IPv6. The migration to IPv6 will also mean that many of the technical obstacles to end-to-end security that are found in IPv4 will likely disappear or be mitigated by the advent of IPv6.

So at the same time as the number of IP-addressable network elements is expected to rise sharply it will also be possible to introduce true end-to-end communication. This will theoretically result in the need for authentication between all IP-addressable network entities for all UMTS operators. This amounts to $(n \cdot (n - 1))/2$ mutual authentication associations, where n is the number of all IP-addressable UMTS network entities. The need for a truly scaleable authentication framework will therefore be strong at that time.

With this as the background, SA3 has started work to prepare for the introduction of a Public Key Infrastructure (PKI) to support *digital certificates* in the core network. These digital certificates will be used solely for authentication purposes for the core network elements. IPsec/IKE has been designed with support for various authentication methods and IPsec/IKE can use digital certificates for authentication. This means that IPsec and IKE can still be used for provision of security services. Furthermore, IPsec/IKE is sufficiently flexible to allow for gradual migration to PKI and digital certificates as the preferred authentication method.

4.2 Authentication Framework

A new *work item* is being created in SA3 to address this issue. The new work item, which will likely result in a new technical specification, has tentatively got the title “*Network Domain Security; Authentication Framework (NDS/AF)*”. The main purpose of NDS/AF is to provide PKI based entity authentication for network elements participating in NDS/IP secured communication.

The introduction of PKI in NDS/IP will probably be executed in phases. For the first phases, the requirement to always use chained tunnels will likely be kept. This will facilitate smooth migration towards full PKI support in NDS/IP for all operators. For instance, an operator may choose to introduce PKI very early within its own security domain while still supporting the use of pre-shared secret towards external destinations. Figure 3 shows a case where PKI based authentication is used for IKE phase-1 over the Za-interface between the SEGs. In this case, operator A may also use PKI internally while operator B uses pre-shared secrets internally.

When PKI is fully integrated in NDS/IP and when IPv6 is widely deployed in the UMTS core networks, the requirement in NDS/IP to use a chained-tunnel architecture will likely be relaxed to allow for direct end-to-end communication between NDS/IP peers.

Note that there is no strong need for PKI for MAPsec since the number of MAPsec security associations is limited by the number of security domains and not by the number of MAP network elements. Therefore, the use of pre-shared secrets for authentication of the MAPsec IKE phase-1 session is unlikely to introduce scalability problems for the foreseen lifetime of MAPsec⁴⁾.

5 Summary

During the last two years the development of Network Domain Security in UMTS has come a long way.

Through the standards defined in 3GPP TS 33.200 and 3GPP TS 33.210 it will be possible to security protect the MAP protocol and GTP-C and other IP-based protocols in the UMTS core network control plane.

The work to refine the security protection in the core network will continue. Automated key management will be developed for MAPsec in order to facilitate robust interworking and management for MAP security. Improved authentication services based on PKI are planned for the NDS/IP. Support for new transport layer protocols like SCTP as well as support for new cryptoalgorithms like AES will be added when they become available from the IETF.

A remaining obstacle to security in the UMTS core network control plane is for operators to actually deploy the security protocols in their networks. Let us hope that operators will adopt the new security standards and use the security services aggressively.

6 Acronyms and Abbreviations

The technical report 3GPP TR 21.905 [16] contains all the official acronyms and abbreviations that apply to more than one technical specification group. In addition, most technical specifications and reports contain a list of abbreviations with relevance to the respective document.

For the purposes of the present document, the following abbreviations apply:

3GPP	3rd Generation Partnership Project (www.3gpp.org)
AES	Advanced Encryption Standard
CR	Change Request
DoI	Domain of Interpretation
f6	MAP encryption algorithm
f7	MAP integrity algorithm
IETF	Internet Engineering Task Force (www.ietf.org)
IKE	Internet Key Exchange
IP	Internet Protocol
IPsec	IP security protocols (as defined in RFC 2401)
ISAKMP	Internet Security Association and Key Management Protocol
KAC	Key Administration Centre
MAC	Message Authentication Code
MAP	Mobile Application Part
MAP-NE	MAP Network Element
MAPsec	MAP security – the MAP security protocol suite
NDS	Network Domain Security
NDS/IP	Network Domain Security; IP network layer security
NDS/MAP	Network Domain Security; MAP application layer security
NE	Network Entity or Network Element
NIST	National Institute of Standards and Technology (www.nist.gov)
PKI	Public Key Infrastructure
RFC	Request For Comment (IETF standards are published as standards track RFCs)
SCTP	Stream Control Transmission Protocol
SEG	Security Gateway
SA	Security Association
SA3	Services and system Architecture: Work Group 3 (Security)
SS7	Signalling System No. 7
TCP	Transmission Control Protocol
UDP	User Datagram Protocol
UMTS	Universal Mobile Telecommunications System

⁴⁾ Observe that MAP itself may live longer than MAPsec. MAP may today be run over the IP protocol stack. NDS/IP could then conceivably be used to secure MAP instead of MAPsec. On the other hand, MAPsec, which is an application layer solution, can be used for both MAP-over-SS7 and MAP-over-IP.

Appendix A Resources

3GPP Resources

All meeting contributions, technical reports and technical specifications are found at the 3GPP web and ftp sites. The web site is at www.3gpp.org and the ftp site is at [ftp.3gpp.org](ftp://ftp.3gpp.org).

On the website under the “specification” title one will find information about the current set of specifications, the latest approved versions, etc. This is very convenient. The 3GPP website also contains a lot of other useful information and is well worth a closer look if one is interested in UMTS standardization.

IETF Resources

A number of 3GPP specifications draw on IETF standards. The IETF standards can be found at the www.ietf.org website.

The email Exploder Lists

For those interested in following the email communications of the technical specification groups (TSGs) there are essentially two ways to do this.

The first is to subscribe to the list(s) of interest. This is done at the at list server webpage at <http://list.3gpp.org/>. You simply pick out the groups of interest to you and subscribe to them. Be warned that many of the groups have a high volume of email contributions to the lists.

The second is simply to use the list server web pages to browse the up-to-date archives of all the listserver email lists. This is convenient since one does not have to subscribe to a list to be able to browse it.

7 References

- 1 Kjøien, G M. Overview of UMTS security for Release 99. *Teletronikk*, 96 (1), 102–107, 2000.
- 2 3GPP. *Mobile Application Part (MAP)*. (3GPP TS 29.002)
- 3a 3GPP. *Network Domain Security; MAP application layer security (Release 4)*. (3GPP TS 33.200)
- 3b 3GPP. *Network Domain Security; MAP application layer security (Release 5)* (3GPP TS 33.200)
- 4 3GPP. *Network Domain Security; IP network layer security (Release 5)*. Draft. (3GPP TS 33.210)
- 5 IETF. *The MAP Security Domain of Interpretation for ISAKMP*. (Work in progress.) <http://www.ietf.org/internet-drafts/draft-arkko-map-doi-04.txt> (Note: Internet-Drafts are removed after 6 months.)
- 6 3GPP. *Tunnelling Protocol (GTP) across the Gn and Gp interfaces*. (3GPP TS 29.060 GPRS)
- 7 IETF. *Security Architecture for the Internet Protocol*. (RFC-2401)
- 8 IETF. *IP Authentication Header*. (RFC-2402)
- 9 IETF. *IP Encapsulating Security Payload (ESP)*. (RFC-2406)
- 10 IETF. *The Internet Key Exchange (IKE)*. (RFC-2409)
- 11 IETF. *User Datagram Protocol*. (RFC-768)
- 12 IETF. *Transmission Control Protocol*. (RFC-793)
- 13 IETF. *Stream Control Transmission Protocol*. (RFC-2960)
- 14 NIST. *Specification for the Advanced Encryption Standard (AES)*. (FIPS-197) (Copies can be obtained at <http://csrc.nist.gov/encryption/aes/>)
- 15 3GPP. *Using PKI to provide network domain security (Telenor/Nokia)*. (Temporary document S3-010622.)
- 16 3GPP. *Vocabulary for 3GPP specifications*. (3GPP TR 21.905)