Low-Delay Sensing and Transmission

JOHANNES KRON

KTH Electrical Engineering

Doctoral Thesis
Stockholm, Sweden 2011

KTH Electrical Engineering

Low-Delay Sensing and Transmission

JOHANNES KRON

Doctoral Thesis in Telecommunications
Stockholm, Sweden 2011

Low-Delay Sensing and Transmission

TRITA-EE 2011:039
ISSN 1653-5146
ISBN 978-91-7415-983-7

Communication Theory
School of Electrical Engineering
KTH (Royal Institute of Technology)
SE-100 44 Stockholm, Sweden

Printed by Universitetsservice US-AB.

To Ulrika

Abstract

This thesis studies cooperative sensing and transmission in the context of wireless sensor networks (WSNs). We especially focus on two means of cooperative sensing and transmission, namely, distributed source coding and relaying. We consider systems where the usefulness of the measured data is dependent on how old the data is and we therefore need *low-delay* transmission schemes. At first sight, the low-delay criterion may seem to be of little relevance, but it is this aspect in particular that distinguishes this thesis from many of the existing communication theoretic results, which often are asymptotic in the block lengths. The thesis is composed of an introductory part, discussing the fundamentals of communication theory and how these are related to the requirements of WSNs, followed by a part where the results of the thesis are reported in Papers A–H.

Papers A–D study different scenarios for distributed source–channel coding. In Paper A, we consider transmission of correlated continuous sources and propose an iterative algorithm for designing simple and energy-efficient sensor nodes. In particular the cases of the binary symmetric channel as well as the additive white Gaussian noise channel are studied. In Paper B, the work is extended to channels with interference and it is shown that also in this case there can be significant power savings by performing a joint optimization of the system. Papers C and D use a more structured approach and propose side-information-aware source–channel coding strategies using lattices and sinusoids.

In Paper E, we apply the methods we have used in joint source–channel coding to the famous Witsenhausen counterexample. By using a relatively simple iterative algorithm, we are able to demonstrate the best numerical performance known to date.

For the case of systems with relays, we study the transmission of a continuous Gaussian source and the transmission of an uniformly distributed discrete source. In both situations, we propose algorithms to design low-delay source–channel and relay mappings. By studying the structure of the optimized source–channel and relay mappings, we provide useful insights into how the optimized systems work. These results are reported in Papers F and G.

In Paper H, we finally consider sum-MSE minimization for the Gaussian multiple-input, multiple-output broadcast channel. By using recently discovered properties of this problem, we derive a closed-form expression for the optimal power allocation in the two-user scenario and propose a conceptually simple and efficient algorithm that handles an arbitrary number of users.

Throughout the thesis we show that there are significant gains if the parts of the system are jointly optimized for the source and channel statistics. All methods that are considered in this thesis yield very low coding and decoding delays. In general, nonlinear mappings outperform linear mappings for problems where there is side-information available. Another contribution of this thesis is visualization of numerically optimized systems that can be used as inspiration when structured low-delay systems are designed.

Keywords: Cooperative communication, wireless sensor networks, low-delay transmission, joint source–channel coding, distributed source coding, estimation, quantization.

Acknowledgments

Five years of challenging work have finally come to an end. During these years as a PhD student there are several people to whom I am grateful.

First and foremost, I am indebted to Prof. Mikael Skoglund without whom this thesis would not have been written. Mikael has been truly generous in every way and has continuously supported me in my research — always prepared to clarify the intriguing details in information theory.

I thank Prof. Fady Alajaji for welcoming me as a visiting research student in his group at Queen's University, Canada, during the fall 2010. The stay was a great experience for me and I appreciate both the research collaboration and your hospitality. I am also thankful to Dr. Niklas Wernersson, Dr. Daniel Persson, Prof. Erik G. Larsson, Dr. Ather Gattami, and Dr. Tobias J. Oechtering for joint publications — it has been a privilege working with you.

I have enjoyed the friendly and stimulating research atmosphere at "plan 4" — thanks to everyone! Some people that deserve special mentioning are: Majid and Hieu for being great roommates, Mattias for helping out with the abstract thinking, Sha for taking beautiful wedding pictures, Ragnar for teaching me a thing or two about bass playing, Annika and Iréne for taking care of administrative issues, the computer support team for making sure the computers run smoothly, Hieu and Nicolas for proofreading the introduction of the thesis, and the Q Restaurant for serving my daily nutritious meals.

I thank Prof. Tor A. Ramstad for taking time to act as opponent for this thesis and also the grading committee for their involvement in the thesis defence.

I want to express my deepest gratitude to my family for always supporting me and for encouraging my technical interests; my parents deserve special thanks for their devotion. Ulrika, you color my life, I am overwhelmingly glad that you are my lovely wife. At last, thanks to my heavenly Father for the hope you have given me.

Johannes Kron
Stockholm, April 2011

Contents

Part I

Introduction

Introduction

We are living in a time of information. Not many of us could have foreseen the rapid progress during the last 20–30 years. Using today's smartphones we have the collected information of the world at arm's length. A strongly contributing cause of this development is the birth of information theory [1] with the advances in telecommunications that have followed. In the beginning of this revolution, much of the research was on point-to-point communication systems. However, in the last years the focus has shifted more and more towards networked systems. In this thesis we are particularly interested in one kind of networks, namely, wireless sensor networks (WSNs) that are expected to play an important role in the future society.

Two important concepts in our analysis of communication strategies for WSNs are source coding and channel coding. Source coding concerns finding good representations for the information that is to be transmitted and channel coding is used to protect the data during the transmission. Source codes often remove redundancy from the source signal, such that it can be represented with less amount of data; an excellent example of an efficient source code is mp3, which is used to represent audio signals. In contrast, channel codes add redundancy to the signal such that if an error occurs, it can either be detected or corrected; an example of this can be found on CDs (in this case the CD can be seen as the channel), which in many cases can be played without any audible errors even if there are scratches on the surface. Source and channel coding are often treated independently and a fundamental result in information theory states that there is no loss of optimality in doing so. But there is a snag in this result: the source and channel codes are assumed to be of infinite lengths. Codes of infinite lengths also give infinite coding and decoding delays. In this thesis we are interested in communication subject to *low-delay* constraints and the assumptions of infinite lengths are clearly infeasible. Therefore, we exclusively study joint source–channel coding schemes, which give extremely low delays and in many cases a significant gain over a separate design. It is the low-dimensional and joint treatment of source and channel coding that differentiate the results in this thesis from many of the already existing results in communication theory.

Outline

During the years as a PhD student, I have been involved in the work with the following papers[1] [2–16] of which [2–9] are formally included in the thesis in Part II (labeled Paper A–H). The remaining of Part I is organized as follows: In Section 1, we introduce the concept of WSNs and give examples of some applications. In Section 2–5, we look at the fundamentals of communication theory and relate these to the requirements of WSNs. This leads to Section 6, where we present the contributions of the thesis, and Section 7, where we draw conclusions and present directions for future work.

1 Wireless Sensor Networks

The emerging technological field of WSNs is a multidisciplinary area involving several research fields, such as electrical circuits design, computer science, signal processing, and communication theory [17–19]. It is the joint advances in these areas that have made the progress in WSNs possible.

A WSN typically consist of small, inexpensive, battery-powered nodes with sensing and communication capabilities. Due to the excessive number of sensors and their sometimes inaccessible locations, it is often desirable that the battery lasts for the entire lifetime of the sensor node, which ranges from a couple of days to several years depending on the specific application. This puts high demands on the design of WSNs in the sense that the nodes need to be highly energy efficient and spend as little energy as possible on signal processing and communication. The sensor nodes are usually spatially distributed and observe some physical phenomenon that evolves over time. One key idea in the design of WSNs is that even though each sensor has limited sensing and communication capabilities on its own, the joint effort of the WSN should provide reliable sensing. This is accomplished by the means of distributed estimation, which takes the tempo-spatial correlation of the observed phenomenon into account, and multi-hop communication schemes that extend the otherwise limited communication ranges.

One can consider two potential scenarios for WSNs. In this thesis, we consider a system that continuously makes decisions and acts based on the measured data. In this scenario, the usefulness of the measured data is dependent on how old the data is and we therefore need *low-delay* transmission schemes. What is to be considered as low delay is relative to the dynamical nature of the physical process that is being observed and the sample rate at which the process is sampled. For example, let us assume that we observe a slowly varying process and take measurements at a sample rate of 0.01 Hz; a low-delay constraint in this example simply means

[1] The author was formerly known as Johannes Karlsson.

that we cannot collect several measurements and transmit these jointly, but we must transmit each single measurement one at a time. The contrasting scenario is a passive system where the data is collected and purely used for offline analysis with no strict delay requirements. In this case we can collect several measurements and transmit blocks of measurements rather than single measurements. The distinction between these two scenarios is necessary because many of the results in information and communication theory rely on coding schemes with infinite delays, which of course is infeasible in the low-delay scenario that we focus on in this thesis. Whether there is a low-delay constraint or not depends on the specific application. Even though the application does not require low-delay transmission, it is sometimes useful to consider low-dimensional systems anyhow due to their low computational complexity.

Applications

WSNs have been proposed to be used in many applications, such as environmental monitoring, industrial automation, intelligent building control, localization, structural health monitoring, surveillance, vehicle-to-vehicle communication, and so on [18, 19]. Since this is a new and emerging field, new applications are continuously being discovered as the technology develops. In the following, we shall elaborate a bit more on a few selected applications.

Intelligent Building Control

Buildings are responsible for 20–40 % of the total energy consumption in developed countries [20]. To meet the high climate goals that are set in terms of reduced carbon dioxide emissions, it is clear that the energy consumption in buildings must be reduced. Some scientists even go further and suggest zero net energy buildings [21], that is, buildings that produce energy (in a sustainable way) according to their own needs and in this way consume zero net energy from the public power plants. One important contributing element for this to become reality is to reduce the overall energy consumption of the building such that the amount of energy that needs to be locally produced is minimized. ABB, a global power and automation company, states the following natural design rules [22]: (1) only use energy when it is really required, (2) only use the amount of energy actually required, and (3) apply the energy that is used with the highest possible efficiency. To accomplish these goals it is crucial to monitor the building and keep track of temperature, oxygen level, number of people, and so on, in each room. WSNs are expected to have a key role in this evolving progress. Other examples of uses for WSNs in buildings are access control and intrusion detection.

Industrial Automation

Industrial automation systems is another application where WSNs are expected to have a big influence [23]. Traditionally, industrial automation systems are realized through wired communications. Sensors are used in industrial processes to monitor critical parameters such that proper counteractions can be initiated immediately when any deviation from the normal state is observed. Some of the benefits of WSNs are ease of installation, maintenance, and service. On the other hand, there are several technical challenges involved in designing a WSN for industrial automation; for example, the environment is harsh with fading, metal obstacles, and radio-frequency interference from machines. There are also strict quality-of-service requirements in terms of accuracy and low communication delays. If the low-delay constraints are not fulfilled, it may even happen that a process becomes unstable.

Structural Health Monitoring

Bridges and other big structures are regularly inspected due to governmental safety regulations. The inspections are time consuming and require sophisticated tools, which are usually expensive and bulky [19]. The idea behind structural health monitoring using WSNs is to deploy a large number of sensor nodes on the structure and measure the response to external excitations. The excitation could either be ambient, due to earthquakes or strong wind, or forced, due to an impact hammer. By analyzing the correlations between the different sensors' measurements it is possible to identify damages such as corrosion and cracks. In [24], it is reported about a prototype that was deployed at the Golden Bridge in San Francisco.

2 Channel Coding

The nodes in a WSN communicate with each other over wireless channels. Before we look at some of the special properties of wireless channels, we will look at some simpler channels. A point-to-point communication channel (see Figure 1) can be described by a conditional probability density function (pdf) of the form $p(\boldsymbol{y}|\boldsymbol{x})$, where $\boldsymbol{x} = (x_1, \ldots, x_N)$ is the transmitted signal vector and $\boldsymbol{y} = (y_1, \ldots, y_N)$ is the received signal vector at the destination. A channel is memoryless and time invariant if the conditional pdf can be factorized according to $p(\boldsymbol{y}|\boldsymbol{x}) = \prod_{i=1}^{N} p(y_i|x_i)$. All channels that are considered in this thesis are assumed to be memoryless and time invariant.

Two simple models for communication channels are the binary symmetric channel (BSC) and the additive white Gaussian noise channel (AWGN). The BSC is a discrete-amplitude (or digital) channel, where the signals that

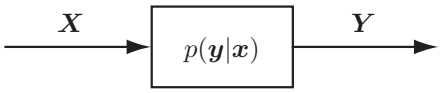


Figure 1: A point-to-point communication channel.

are transmitted and received are binary, that is, $x, y \in \{0, 1\}$. The conditional probability mass function[2] (pmf) that describes the relation between the input and the output is given by

$$P_{\text{BSC}}(y|x) = \begin{cases} 1-\epsilon, & \text{if } y = x, \\ \epsilon, & \text{if } y \neq x. \end{cases} \tag{1}$$

The channel can be modeled as someone rolling a dice for each transmitted symbol and depending on the outcome, the transmitted signal reaches the destination unchanged or with an error. For example, the bit could be changed if the outcome is one and unchanged for all other outcomes. In this example, the probability of a bit error is $\epsilon = 1/6$. For the AWGN channel, the input and output symbols are real numbers, $x, y \in \mathbb{R}$, it is therefore referred to as a continuous-amplitude (or analog) channel. The relation between x and y is such that $y = x + n$, where n is white Gaussian noise with variance σ_N^2. The channel's conditional pdf is given by

$$p_{\text{AWGN}}(y|x) = \frac{1}{\sqrt{2\pi\sigma_N^2}} \exp\Big(- \frac{(y-x)^2}{2\sigma_N^2} \Big). \tag{2}$$

As can be seen from (1) and (2), there is some uncertainty about what was transmitted when using these channels. For example, if we receive 1 on the BSC we cannot be sure if 0 or 1 was transmitted. Channel coding is what makes it possible to reliably transmit important information, such as phone calls, text messages, business documents, measurements, and so on, over channels that, by themselves, are unreliable. The basic idea of channel coding is to transmit blocks of data rather than single symbols. This is done by taking k information symbols and mapping them to a vector $\boldsymbol{x}$ of length N, which we transmit on the channel. The rate at which we transmit information is defined as $R \triangleq k/N$. Shannon showed [1] that, if the block length N approaches infinity and as long as R is less than a number denoted C, it is possible to transmit reliably without any errors. He further showed that, if the rate is higher than C, the probability of error is bounded away from zero. The number C is called the capacity of the channel. For discrete memoryless channels, the capacity is given by [25, Ch. 7]

$$C = \max_{P(x)} I(X;Y), \tag{3}$$

[2]The pdf is denoted with $p(x)$ and used for continuous random variables, whereas the pmf is denoted $P(x)$ and used for discrete random variables.

where $I(X;Y)$ is the mutual information between X and Y, and the maximum is taken over all possible input distributions $P(x)$. The capacity of a continuous channel can be computed in a similar way, but in this case we need a constraint on the transmission power. The capacities of the BSC and the AWGN channel are given by [25, Chs. 7 & 9]

$$C_{\text{BSC}} = 1 + \epsilon \log_2(\epsilon) + (1-\epsilon)\log_2(1-\epsilon) \quad \text{bits per channel use,} \tag{4}$$

$$C_{\text{AWGN}} = \frac{1}{2}\log_2\left(1 + \frac{P}{\sigma_N^2}\right) \quad \text{bits per channel use,} \tag{5}$$

respectively, where P is the transmission power in the case of the AWGN channel. The signal-to-noise ratio (SNR) of the AWGN channel is defined by $\text{SNR} = P/\sigma_N^2$.

If it is possible to get arbitrarily low error probability at a certain rate R, via coding, that rate is said to be *achievable*. Shannon's proof that rates below C, but not above, are achievable was theoretical and did not include any practical guidelines on how to construct a channel code that achieves the capacity. It is not until much later, with the advent of turbo codes [26] and the rediscovery of LDPC codes [27, 28], that communication close to the capacity has been made possible. A channel code enforces a structure on the transmitted data, making the data robust against the channel noise. Returning to the dice example, if you roll a dice one time, you cannot be sure of the outcome. However, if you roll the dice 6000000 times, the number of ones will approximately be 1000000. In a similar way, the effect of the channel noise becomes predictable for codes with long block lengths.

Wireless Channels

Two characteristic properties of a wireless channel are its broadcasting nature and interference between users. In the most general form, if K users are transmitting and receiving at the same time, this can be modeled by a conditional pdf like $p(\boldsymbol{y}_1, \ldots, \boldsymbol{y}_K | \boldsymbol{x}_1, \ldots, \boldsymbol{x}_K)$, where user i transmits $\boldsymbol{x}_i$ and receives $\boldsymbol{y}_i$. A special case would be if only one user transmits, in which case we have a broadcast channel, or if only one user receives, in which case we a multiple-access channel. These two cases correspond to $p(\boldsymbol{y}_1, \ldots, \boldsymbol{y}_K | \boldsymbol{x})$ and $p(\boldsymbol{y} | \boldsymbol{x}_1, \ldots, \boldsymbol{x}_K)$, respectively. There are different ways to deal with interference. The simplest way is by using orthogonal access techniques, such as time-division multiple access (TDMA) and frequency-division multiple access (FDMA), where the users transmit in different time slots or frequencies, respectively. In general, there can be some loss in using orthogonal techniques in the sense that the entire capacity region[3] cannot be achieved.

[3]The multi-user equivalence of capacity as discussed earlier is a capacity region $\mathcal{C}$, where for example in a two-user scenario, a rate pair (R_1, R_2) is achievable if and only if $(R_1, R_2) \in \mathcal{C}$.

Despite this, the simplicity of the orthogonal techniques makes them useful in practical systems [25, Ch. 15].

Another fundamental property of wireless channels is fading [29]. Fading is a way to describe how the quality of the channel varies with time and frequency. There are two kinds of fading, large-scale fading and small-scale fading. Large-scale fading is due to path loss as a function of distance and shadowing by large objects. Small-scale fading is due to constructive and destructive interference caused by multi-path propagation. Throughout this thesis, we assume that the variations are slow enough so that the channel quality remains constant for the entire transmission. In most cases where we optimize the transmission scheme, we assume the instantaneous channel quality to be known by all nodes. If the channel quality deviates from the assumed channel quality, we have a mismatch condition in which there is some loss in terms of performance.

3 Source Coding

The sensor nodes measure physical quantities, such as temperature, oxygen level, humidity, pressure, vibrations, etc. In communication theory, this measured quantity is referred to as the source. In this thesis, we are interested in discrete-time sources. A discrete-time source $\{X_i\}_{i=-\infty}^{\infty}$ can be viewed as an indexed sequence of random variables (RVs). The dependencies between the RVs could be arbitrary and is specified by a joint pdf, $p(\ldots, x_{i-1}, x_i, x_{i+1}, \ldots)$. An important special case is when the RVs are independent and identically distributed (i.i.d.), in which case the pdf can be factorized as

$$p(x_k, \ldots, x_{k+K}) = \prod_{i=k}^{k+K} p(x_i), \tag{6}$$

for all $k, K \in \mathbb{Z}$. For i.i.d. sources we sometimes use a simplified notation and let the RV (without indices) denote the source, that is, the i.i.d. source $\{X_i\}_{i=-\infty}^{\infty}$ is denoted simply by X. A RV could either be discrete amplitude, which means that it can take a finite number of possible values, or continuous amplitude, in which case it could take any value out of an infinite set of values. Source coding deals with finding a good representation (typically binary) for all these sources so that they can be stored or transmitted. There is a distinction between two kinds of source coding — lossless and lossy.

3.1 Lossless Source Coding

In lossless source coding, the representation is such that the source can be perfectly reconstructed. This is, for example, the way source coding of

text is done. The efficiency of a source code is measured by the number of bits/sample needed to represent a source. The same source can have different representations, where the most efficient representation is the one that uses the least number of bits. Shannon showed [1] that there exists a lower bound on the number of bits/sample needed to represent a specific source, so that it can be reconstructed without any errors. The lower bound is called the entropy rate of the source and measures the uncertainty of the source. If the source is i.i.d., the entropy rate of the source is the same as the entropy of the RVs that constitute the source. The entropy of a discrete RV is denoted $H(X)$ and defined by [25, Ch. 2]

$$H(X) \triangleq -\sum_{x} P(x) \log_2(P(x)), \tag{7}$$

where $P(x)$ is the pmf of X. To achieve the entropy rate when coding a source, it is generally required that the source is encoded in blocks of length that tends to infinity. Two entropy-achieving coding strategies are Huffman [30] and Lempel–Ziv [31,32] coding. Huffman coding has the advantage of working for short blocks of data but requires the probability distribution to be known, whereas Lempel–Ziv coding makes no assumption other than that the source is stationary and ergodic.

3.2 Lossy Source Coding

For continuous sources, such as measurements in WSNs or audio signals, it is not possible to find a representation such that perfect reconstruction is possible; an infinite number of bits would be needed for this to be possible. This problem is solved by introducing lossy source coding. In this case we do not require the reconstruction $\{\hat{X}_i\}$ to be perfect but we are satisfied as long as the reconstruction is close to the original source signal $\{X_i\}$. Two lossy source codes that are widely used are JPEG for digital images and MPEG-1 layer 3 (mp3 in short) for audio signals. A measure of the closeness to the original signal is needed in order to evaluate the performance of different lossy source coding schemes. It is common to use a measure that works on a symbol-by-symbol basis. The measure, denoted $d(x, \hat{x})$, should be such that $d(x, \hat{x}) \geq 0$, with equality if and only if $x = \hat{x}$. One measure that satisfies this general condition is the squared-error distortion, where

$$d(x, \hat{x}) = (x - \hat{x})^2. \tag{8}$$

The measure can be generalized to sequences by averaging over the sequence in the following way:

$$d\left(\{x_i\}_{i=1}^N, \{\hat{x}_i\}_{i=1}^N\right) = \frac{1}{N} \sum_{i=1}^{N} d(x_i, \hat{x}_i). \tag{9}$$

By taking the expected value of $d\left(\{X_i\}_{i=1}^N, \{\hat{X}_i\}_{i=1}^N\right)$ we get the mean squared error (MSE), which is by far the most widely used distortion measure. One explanation for this is that it makes the analysis tractable in many situations. Further motivation for using the MSE can be found in estimation theory, where the minimum MSE estimate minimizes the distortion among a large class of distortion measures [33].

In the lossless case, we saw how the entropy rate determined a lower bound on how many bits that are needed to represent a source. For the lossy case, Shannon (once again) showed [1] that there is a minimum rate $R(D)$ (bits/sample) that is required to achieve a certain distortion D. This relationship is characterized in rate distortion theory [25, Ch. 10]; the rate distortion function for an i.i.d. source X is given by

$$R(D) = \min_{p(\hat{x}|x):\mathbf{E}[d(X,\hat{X})]\leq D} I(X;\hat{X}), \tag{10}$$

where $I(X;Y)$ denotes the mutual information between the RVs X and Y. Evaluated for a memoryless Gaussian source with variance σ_X^2 and MSE as distortion measure, the rate distortion function is given by

$$R(D) = \begin{cases} \dfrac{1}{2}\log_2 \dfrac{\sigma_X^2}{D}, & \text{if } D < \sigma_X^2, \\ \\ 0, & \text{if } D \geq \sigma_X^2. \end{cases} \tag{11}$$

Alternatively, we can rewrite this expression to obtain the distortion as a function of the rate $R \geq 0$:

$$D(R) = \sigma_X^2 2^{-2R}. \tag{12}$$

For correlated sources it is possible to achieve an even lower distortion. For example, the distortion rate function for an i.i.d. Gaussian source $\boldsymbol{X} = (X_1, X_2)^T$ with covariance matrix

$$\mathbf{E}[\boldsymbol{X}\boldsymbol{X}^T] = \begin{pmatrix} \sigma_X^2 & \rho\sigma_X^2 \\ \rho\sigma_X^2 & \sigma_X^2 \end{pmatrix}, \tag{13}$$

where ρ denotes the correlation coefficient, is given by [34, Ch. 4]

$$D(R) = \begin{cases} \sigma_X^2 \dfrac{1 - |\rho| + (1+|\rho|)2^{-4R}}{2}, & \text{if } R \leq \dfrac{1}{4}\log_2 \dfrac{1+|\rho|}{1-|\rho|}, \\ \\ \sigma_X^2 \sqrt{1-\rho^2} 2^{-2R}, & \text{if } R > \dfrac{1}{4}\log_2 \dfrac{1+|\rho|}{1-|\rho|}. \end{cases} \tag{14}$$

The interpretation of the expression in (14) is that if the source $\boldsymbol{X} = (X_1, X_2)^T$ is encoded with R bits/sample, then $D(R)$ is a lower bound

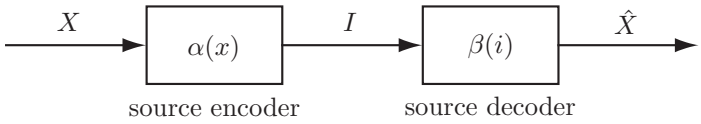


Figure 2: Source coding by scalar quantization.

on the distortion, that is, $\frac{1}{2}\mathbf{E}[\|\boldsymbol{X} - \hat{\boldsymbol{X}}\|^2] \geq D(R)$ for any source code. By comparing (12) and (14) and assuming that the correlation $|\rho| < 1$, we can see that gain from the correlation is $\sqrt{1-\rho^2}$ at high rates. Once again, the bounds given by (12) and (14) are asymptotic, meaning that they are in general achievable only if an infinite number of source samples are encoded jointly.

3.3 Quantization

A simple and practical lossy source coding method is scalar quantization (SQ), see Figure 2. In this method, a continuous-amplitude variable $x \in \mathbb{R}$ is first mapped to an index $i \in \{1, \ldots, M\}$ by a source encoder α according to

$$i = \alpha(x) \quad \text{if } x \in \Omega_i. \tag{15}$$

The mapping is determined by the sets $\{\Omega_i\}_{i=1}^M$, which partition the real line into disjoint quantization regions; the number M is typically a power of two, $M = 2^b$, which means that the index can be represented by a sequence of b bits that could be stored or transmitted. The source decoder β uses a reconstruction codebook $\mathcal{W} = \{w_1, \ldots, w_M\}$ to form an estimate of x from the index,

$$\hat{x} = \beta(i) = w_i. \tag{16}$$

The simplest form of SQ is uniform quantization, where all quantization regions Ω_i (except the endpoints) are intervals of the same length. For low bit rates, it is generally beneficial to optimize the quantization regions for the source distribution. Lloyd [35,36] and Max [37] have independently developed similar algorithms for designing a scalar quantizer that minimizes the MSE for a given source, the optimized quantizer is commonly known as a Lloyd–Max quantizer. The optimization algorithm consists of the following steps: First choose some initial reconstruction points w_i, $i = 1, \ldots, M$. Given the reconstruction points it will be possible to find the sets $\{\Omega_i\}_{i=1}^M$ that partition the real line such that the MSE is minimized. The condition for optimality can be expressed as

$$\Omega_i = \{x : (x - w_i)^2 < (x - w_k)^2, \forall k \neq i\}. \tag{17}$$

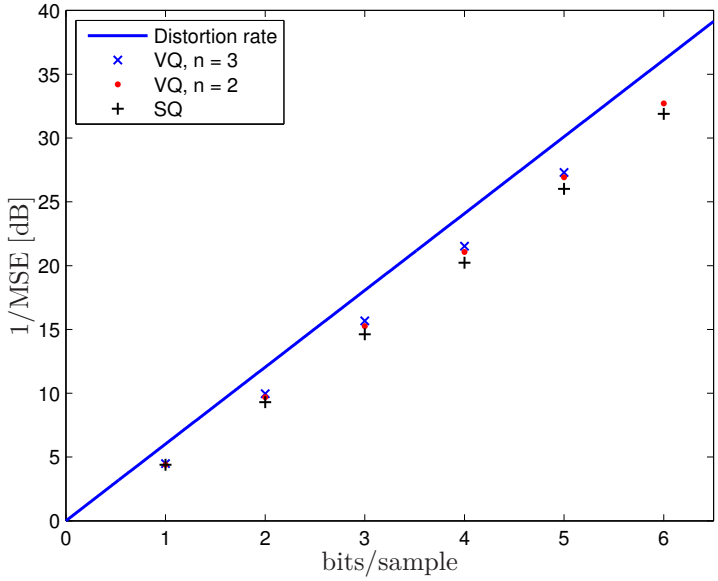


Figure 3: Performance comparison of the distortion rate curve, VQ ($n = \{2, 3\}$), and SQ, for a memoryless Gaussian source.

In a similar fashion, given the sets $\{\Omega_i\}_{i=1}^{M}$ we can find the reconstruction points that minimize the MSE by

$$w_i = \mathbf{E}[X|X \in \Omega_i]. \tag{18}$$

The idea is now to iterate between (17) and (18) until the expressions converge. One drawback of this method is that the final solution will depend on the initial choice of reconstruction points — the algorithm can only be assured to converge to a local minimum.

A straightforward generalization of SQ is vector quantization (VQ). In VQ, instead of mapping a scalar to an index, α operates on a vector $\boldsymbol{x} \in \mathbb{R}^n$, which is mapped to an index i. The Lloyd–Max algorithm can be extended to VQ and is then referred to as the generalized Lloyd algorithm. An important contribution can be found in [38], where the LBG algorithm, which partly solves the initialization problem, is proposed.

One important property of VQ is that it asymptotically (in n) reaches the lower bound on the distortion given by the rate distortion theory. For high rate, the advantages of VQ over SQ can be divided into the following three contributions [39]. When the dimension n increases, the shape of the quantization regions (given by $\{\Omega_i\}_{i=1}^{M}$) will be more and more spherical,

which (if possible) would be their optimal shape — this is called the *space-filling advantage*. If there is correlation between the components in the vector, a VQ can focus on the differences among these, whereas an SQ cannot do this distinction since it operates on each component individually — this is called the *memory advantage*. Even if there is no correlation between the components, there is a *shape advantage* of VQ unless each component is uniformly distributed. For example, samples from a two-dimensional Gaussian source will be spread in a circular pattern around the mean. Hence, there is no need to spend bits on the corner points like a SQ would do. If a memoryless Gaussian source is encoded at high rate, the asymptotic gain (in n) of using VQ over SQ is 4.34 dB — 1.53 dB due to the space-filling advantage and 2.81 dB due to the shape advantage (there is no memory advantage in this case). In Figure 3, the distortion rate curve is plotted together with the performance of VQs of different dimensions. However, there is one major concern with VQ — the complexity grows exponentially with the dimension.

Lattice Quantization

The complexity of VQ can be reduced by enforcing a structure on the quantization regions. Lattice quantization is one example of how this can be done. An N-dimensional lattice Λ is defined by the generator matrix $G \in \mathbb{R}^{N\times N}$. A point $\boldsymbol{l} \in \mathbb{R}^N$ belongs to the lattice if and only if it can be written as $\boldsymbol{l} = G\boldsymbol{i}$, where $\boldsymbol{i} \in \mathbb{Z}^N$ and $\mathbb{Z} = \{0, \pm 1, \pm 2, \ldots\}$. The nearest neighbor quantizer of a lattice Λ is defined by

$$Q_\Lambda(\boldsymbol{x}) \triangleq \arg\min_{\boldsymbol{l}\in\Lambda} \|\boldsymbol{x} - \boldsymbol{l}\|. \tag{19}$$

For many widely used lattices, there exist highly efficient algorithms that perform the operation in (19) without the need of an exhaustive search. For lattices where such methods are not available, there exist closest-point search methods that make use of the lattice structure to significantly improve the performance compared to exhaustive search methods, see for example [40]. However, these methods still have exponential complexity. A comprehensive study of lattices and lattice quantization can be found in [41].

4 Joint Source–Channel Coding

In Section 2 we saw how a given channel was associated with a capacity that defined an upper bound on the rate at which reliable communication is possible. The results in Section 3 further showed that there is a minimum rate needed to represent a source. A fundamental result in information theory states that there is no loss of optimality in doing the source and

channel coding separately, this is called the *source–channel separation theorem* [25, Ch. 7]. However, this theorem heavily rests on the assumptions of infinite block lengths in the source and channel codes. In situations with low-delay constraints, the use of long block lengths is not feasible. Because of this the source–channel separation theorem does not hold anymore and a joint source–channel code can give better performance. One situation where a low delay is crucial is in closed-loop control over wireless channels. In this case a wireless sensor measures some quantity (e.g., temperature) that should be transmitted over a wireless channel to a control system. In the following subsections, different joint source–channel coding strategies that can be used in scenarios like this will be discussed.

4.1 Channel-Optimized Quantization

If the source coding system in Figure 2 is combined with a channel, we get a system as the one shown in Figure 4. In this figure, the sensor node produces an index I for each source sample X. The index is transmitted over the channel and affected by the channel's noise characteristics; because of this the performance will be affected by the index assignment that is used. The index assignment is the way that different quantization regions are assigned to their corresponding indices. In the following, we will focus on discrete channels, where there is a straightforward connection between the quantization indices and the channel input[4]. Different channel symbols will in general have different distance properties. If we, for example, take the BSC with 4 bits per source symbol, the indices 0 and 15 have the binary representations 0000 and 1111, respectively, and can be seen as far apart in the channel space, whereas the index 1, with binary representation 0001, is close to index 0 since only one bit differs. A good index assignment should preserve distance properties, that is, source symbols that are close should be mapped to indices that are close in the channel space and source symbols that are far apart should be mapped to indices that are far apart in the channel space [42]. To find the optimal index assignment, one would have to do an exhaustive search among the $M!$ possible combinations, which is infeasible in most cases. Despite the fact that the optimal index assignment is hard to find, a random index assignment should be avoided. It can even be shown that a random index assignment is asymptotically bad for uniform sources [43]. The performance of different index assignments has been studied in [44].

An alternative approach, to optimizing the index assignment, is to take the channel into consideration in the design algorithm when determining the quantization regions as well as the reconstruction points. This is the

[4]The same reasoning applies to continuous-alphabet channels, but only after a modulation from the discrete indices to the channel space has been defined.

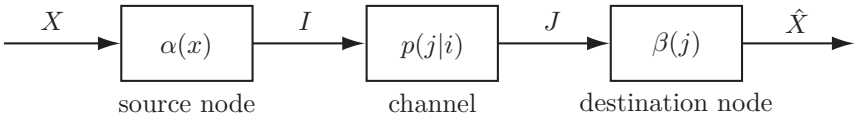


Figure 4: Scalar joint source–channel coding over a digital channel.

strategy of [45], where a generalization of the Lloyd–Max algorithm is used to find channel-optimized quantizers. By taking the effect of the channel into account, the equations corresponding to (17) and (18) are given by

$$\Omega_i = \{x : \mathbf{E}[(x - \hat{X})^2 | I = i] < \mathbf{E}[(x - \hat{X})^2 | I = k], \forall k \neq i\} \tag{20}$$

and

$$c_j = \mathbf{E}[X | J = j], \tag{21}$$

respectively. Similarly to the Lloyd–Max algorithm, the optimized quantizer is found by iterating between applying (20) and (21), updating the quantization regions and the reconstruction points, respectively. By expanding the squares in (20) and rearranging the expression, this update can be done in a very efficient manner [45]. An important observation in [45] is that some Ω_i may be empty, that is, the indices corresponding to these sets will never be transmitted. This will make the system more robust against channel noise at the cost of increasing the quantization distortion. This phenomenon is also observed in Paper A, where a generalization of the design algorithm in [45] is used for distributed source coding. As in the case of SQ, presented in Section 3.3, the optimization algorithm can be generalized to the design of VQs for noisy channels [46–49].

4.2 Bandwidth Compression–Expansion

The physical channels that we use for transmission are analog and not digital. When we talk about digital channels, we implicitly assume that a mapping from our digital channel to the analog physical channel has been provided. The energy efficiency of the communication system could be increased — allowing us to save valuable energy in the sensor nodes — if we mapped source symbols directly to the analog channel. In this section, we will therefore look at analog source–channel mappings, that is, mappings that take a vector of analog values as input and produce a vector of analog values at the output.

Assume that a K-dimensional continuous-amplitude source is to be transmitted over an L-dimensional continuous channel. One could define a modulation from the discrete indices to the continuous channel space,

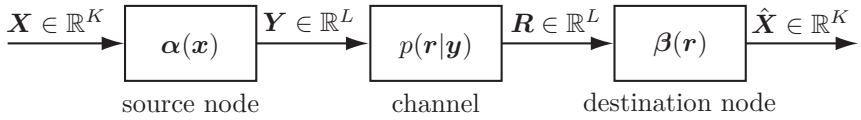


Figure 5: Joint source–channel coding over an analog channel.

and use the strategies that are described in Section 4.1. A more general approach is to let the source encoder map the source symbols directly to the channel space by a mapping $\boldsymbol{\alpha} : \mathbb{R}^K \mapsto \mathbb{R}^L$. At the destination, the decoder estimates the transmitted source symbol by a mapping $\boldsymbol{\beta} : \mathbb{R}^L \mapsto \mathbb{R}^K$. Depending on the ratio K/L, this can be seen as either *bandwidth compression*, for $K/L > 1$, or *bandwidth expansion*, for $K/L < 1$. A system where this approach is employed can be seen in Figure 5. It should be emphasized that, if the jointly optimal pair of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ can be found, this is the optimal strategy for any choice of K and L. This is clear since all other structured communication schemes can be seen as special cases of these arbitrarily mappings (including schemes that use a source code in cascade with a channel code). However, the problem is to find the jointly optimal pair of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, which is not known except for some special cases. One such case is if we have a matched bandwidth between the source and the channel (i.e., $K = L$), the source is memoryless Gaussian, and the channel is an AWGN channel. In this case the optimal strategy is to let $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ be linear functions, that is, to use uncoded transmission. Necessary and sufficient conditions for uncoded transmission to be optimal are given in [50].

The idea of this kind of analog codes was mentioned already by Shannon in [51]. Theoretical characterization of optimal analog communication systems can be found in [52–54]. One important observation is that a linear system is not optimal in general. In the case of bandwidth compression, a linear system would need to discard some of the dimensions of the source. On the other hand for bandwidth expansion, a linear system would only use a K-dimensional subspace of the L-dimensional channel space, which means that it does not use all the available degrees of freedom. For practical results and numerical optimization of α and β, see [55–62] for the case of bandwidth compression and [57,59–64] for the bandwidth expansion case. An example of a 2:1-bandwidth-compression curve (i.e., α) can be seen in Figure 6. The curve has been optimized for a Gaussian source and an AWGN channel using an algorithm similar to the one described in [56]. The two-dimensional Gaussian source samples, $\boldsymbol{x} = (x_1, x_2)$, are mapped to the closest point on the curve that represents the one-dimensional channel space. The upper left end of the curve corresponds to $\alpha(\boldsymbol{x}) = -4$ and the lower right end corresponds to $\alpha(\boldsymbol{x}) = 4$. It is often very hard to find good

pairs of numerically optimized α and β, unless the dimensions are rather low. A solution to this problem is to enforce a structure on α, this is done for the bandwidth expansion case in [65–67]. It is also possible to design hybrid digital and analog systems as done in [68, 69].

5 Cooperative Transmission

In the past years, communication schemes for cooperative transmission have received a lot of attention from the research community. With the increasing popularity and relevance of ad-hoc WSNs, cooperative transmission is more relevant than ever. In the following, we will discuss two important fields in cooperative transmission, namely distributed source coding and relaying.

5.1 Distributed Source Coding

In WSNs, there may be a high correlation between different sensor measurements due to high spatial density of the sensor nodes. This motivates distributed source coding of correlated sources.

Lossless Distributed Source Coding

In Section 3, we saw how the entropy of a discrete source determined the minimum number of bits needed to represent the source without any loss of information. This concept can be generalized to the joint entropy of any number of sources. The joint entropy of the discrete sources X and Y is denoted $H(X,Y)$ and determines the minimum number of bits needed to represent the two sources jointly [25, Ch. 2]. The joint entropy can be divided into two parts,

$$H(X,Y) = H(Y) + H(X|Y), \tag{22}$$

where $H(X|Y)$ is the conditional entropy of X given Y. The conditional entropy determines the minimum number of bits needed to represent X if Y is known to both the encoder and the decoder. A fundamental property of the conditional entropy is $H(X|Y) \leq H(X)$, with equality if and only if X and Y are independent. This means that the number of bits needed to represent the source can be reduced if there is a dependency between X and Y. One remarkable result is the Slepian–Wolf theorem [70], which states that even if only the decoder has access to the side information Y, it is still possible to encode X with only $H(X|Y)$ bits and get a perfect reconstruction at the decoder.

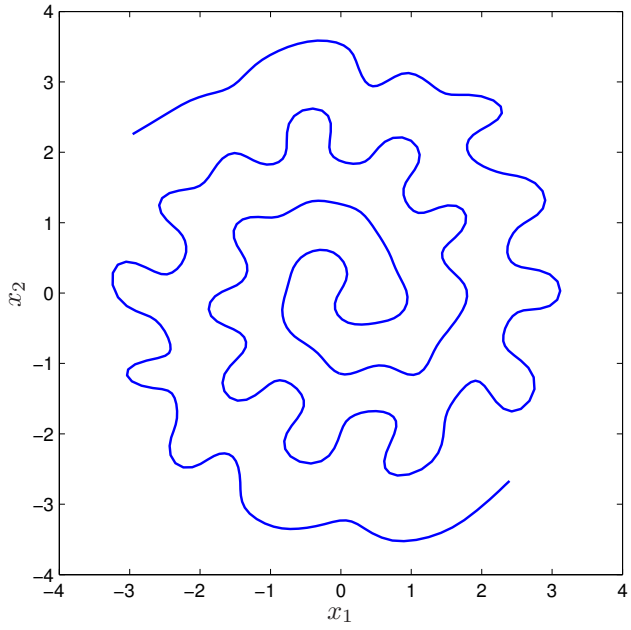


Figure 6: Bandwidth compression curve optimized for a two-dimensional Gaussian source and a one-dimensional AWGN channel with SNR = 30 dB, that is, $K = 2$ and $L = 1$. The two-dimensional Gaussian source samples, $\boldsymbol{x} = (x_1, x_2)$, are mapped to the closest point on the curve that represents the one-dimensional channel space.

Lossy Distributed Source Coding

The results for discrete sources were later extended to the case of lossy source coding of continuous sources in [71] and is then referred to as Wyner–Ziv coding. The distortion rate function in (12) provided the minimum achievable distortion for a memoryless Gaussian source X, encoded with R bits/sample. If we now assume that the decoder has access to side information in terms of another Gaussian source Y that is correlated with X according to

$$\rho = \frac{\mathbf{E}[XY]}{\sqrt{\mathbf{E}[X^2]\mathbf{E}[Y^2]}}, \tag{23}$$

Wyner–Ziv's results state that the distortion rate function with side information at the decoder is given by

$$D^*(R) = D_{X|Y}(R) = \sigma^2_{X|Y} 2^{-2R} = \sigma^2_X (1-\rho^2) 2^{-2R}. \tag{24}$$

Distributed source coding is important in WSNs because it allows us to save energy and bandwidth by reducing the amount of information that needs to be transmitted. However, the results in [70, 71] are nonconstructive in the sense that they rely on random codes of infinite block lengths. Ideas on how to construct practical Slepian–Wolf coding schemes using channel codes are presented in [72–75]. Schemes for lossy distributed source coding can be found in [76, 77]. These schemes are based on long block codes, which introduces a delay in the system. In situations where long delays are a concern, an alternative approach is to look at the problem as a quantization problem; see, for example, [78–81]. In WSNs, the sensor nodes usually run on batteries and operate under strict power constraints. It is therefore relevant to include a noisy channel over which the source coded symbols are to be transmitted. This problem is studied in [82–84].

5.2 Relay Channel

Assume that one node in a wireless network has a message for a distant node. One way to reduce the energy consumption is to let an intermediate node act as a relay and in this way increase the reach of the source node. The three-node relay channel that was introduced in [85] is an example of such scenario. All nodes could in principle act both as transmitting and receiving nodes simultaneously. In this thesis we are interested in the scenario where one node acts as a source node that wants to communicate a message to a destination node. Besides the direct path from the source to the destination, a relay node, with no objectives of its own, assists the communication and creates an alternative path. In theory, the source and the relay may transmit simultaneously for the whole duration of the transmission, but a common

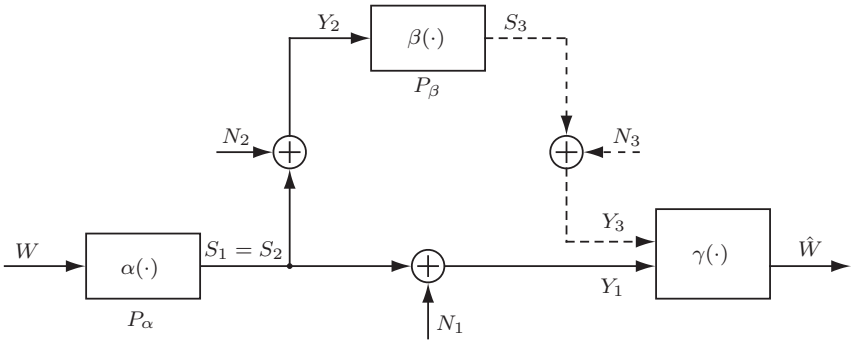


Figure 7: Half-duplex orthogonal AWGN relay channel. Transmissions are subject to the power constraints $\mathbf{E}[\alpha(W)] \leq P_\alpha$ and $\mathbf{E}[\beta(Y_2)] \leq P_\beta$ at the source and relay nodes, respectively.

assumption is that practical limitations of the relay make it unable to receive and transmit simultaneously on the same frequency [86]. Therefore the time is divided into two phases, the first phase in which only the source node transmits and the second phase in which both the source and the relay nodes transmit. Another common assumption, which we adopt in this thesis, is that the source and relay nodes communicate over orthogonal channels (see Figure 7). All in all, this is usually referred to as a half-duplex orthogonal relay channel.

Even though the relay channel has been extensively studied, its capacity is not known in the general case. Some early theoretical results on the relay channel with an upper bound on the capacity and explicit formulas for some degraded cases can be found in [87]. The main problem is to determine how the source and relay nodes should operate. For a fixed strategy it is usually possible to determine achievable rates [88,89]. Some relaying strategies that have been proposed include amplify-and-forward (AF), decode-and-forward (DF), estimate-and-forward (EF), and compress-and-forward (CF). In AF, the relay transmits a linear amplification of what it receives. This works well if the SNR of the channel from the relay to the destination is high. DF requires somewhat more processing by the relay, which in this case decodes the received message and encodes it again before it is transmitted to the destination. In EF, the relay outputs an estimate of the signal that was transmitted from the source, and in CF, the relay uses Wyner–Ziv lossy source coding on the received signal. In some cases, there may be a low-delay constraint on the transmission or a low-complexity operation at the relay may be desirable. One approach in these situations is to let the relay perform a memoryless operation, where the current output only depends on

the current input. This is commonly referred to as *instantaneous relaying*.

6 Contributions of the Thesis

This thesis studies cooperative sensing and transmission in the context of WSNs. By that we mean, using the terminology that was introduced in Sections 2–5, distributed source and channel coding subject to low-delay constraints. At first sight, the low-delay criterion may seem to be of little relevance, but it is this aspect in particular that distinguishes this thesis from many of the existing communication theoretic results, which often are asymptotic in, for example, the block lengths.

In Figure 8, the relation of the included papers are shown in a schematic overview. Papers A–D study different scenarios of distributed source–channel coding. In Paper A and Paper B we use a generalization of the Lloyd–Max algorithm to find optimized systems, whereas in Paper C and Paper D we use a more structured approach by using sinusoids and lattice quantization. In Paper E, we apply the methods that we have found successful in joint source–channel coding to an open problem in the automatic control discipline, namely, Witsenhausen's counterexample. By doing this we are able to demonstrate the best numerical performance known to date. In Paper F and Paper G we turn our attention to the relay channel; new schemes for lossy source–channel coding are considered in Paper F, whereas Paper G deals with designing optimized source–channel mappings for lossless transmission of a discrete source. The optimizations in Papers E–G are also based on a generalization of the Lloyd–Max algorithm. Paper H is an outlier in that it studies how to perform the optimal power allocation for *linear* transmission in a MIMO broadcast channel. In the same way that this is the only paper where we study linear transmission it is also our only result that provides *the* optimal solution in closed form.

We have almost exclusively looked at analog sources and consistently used the MSE distortion measure. This is true for all papers except Paper G, where the source is an information symbol and we use the symbol-error rate as distortion measure. The recurring strategy that we use to accomplish low-delay sensing and transmission is to merge the source and channel coding operations into one single operation as discussed in Section 4. Even though the information theoretic upper bounds on the performance of source and channel coding are often too optimistic when studying low-delay systems, we use these as references since they provide the optimal performance theoretically attainable (OPTA).

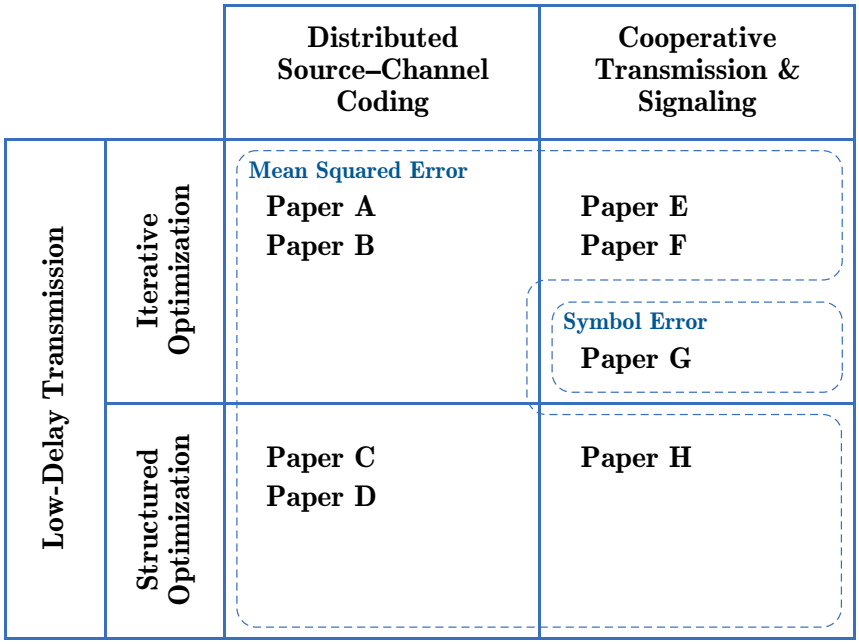


Figure 8: Schematic overview of the included papers.

Summary of Papers A–H[5][6]

Paper A: Distributed Quantization over Noisy Channels [2]

We consider the problem of designing simple and energy-efficient sensor nodes in a WSN. An algorithm for designing distributed scalar quantizers for orthogonal channels is proposed and evaluated. In particular the cases of the BSC as well as the AWGN channel are studied. The system works on a sample by sample basis yielding a very low encoding complexity, at an insignificant delay. Due to the source correlation, the resulting quantizers use the same indices for several separated intervals in order to reduce the quantization distortion. The proposed quantizers can be used in low-complexity transmission schemes for wireless sensor nodes and allow for valuable energy savings.

[5] I changed my name from Johannes Karlsson to Johannes Kron in January 2011.

[6] In Paper A, where I am the second author, I have primarily done the preparations and writing of the part concerning the BSC. In Papers B–H, where I am the first author, I have done the majority of both the preparations and writing.

Paper B: Low-Delay Joint Source–Channel Mappings for the Gaussian MAC [3]

Similarly to Paper A we design simple and energy-efficient sensor nodes. The difference is that we study transmission over an interfering channel instead of orthogonal channels. We propose the use of low-delay joint source–channel mappings and show for the matched bandwidth case how performance saturation, which is unavoidable with linear transmission, can be overcome by optimizing the mappings. To gain deeper understanding of how the optimized systems work, the resulting mappings and their corresponding joint decoders are visualized. The optimized mappings are in general nonlinear and perform a combination of signaling and transmission of analog data. We next look at bandwidth expansion based on 16-QAM and BPSK modulation and show that it is possible to have large performance gains by optimizing the mappings. The implementation aspects of the design algorithm are also discussed in detail.

Paper C: Lattice-Based Source–Channel Coding in WSNs [4]

We consider the problem of gathering measurements in a WSN consisting of a large number of sensor nodes. A practical joint source–channel coding scheme is proposed and evaluated. The scheme uses lattices to extend a previously proposed scheme to higher dimensions. The key idea is to use conventional point-to-point communication for a subset of the sensor nodes and side-information aware transmission for the remaining sensor nodes. We show that, by expanding from one to eight dimensions, a gain of about 1 dB is achievable. The overall transmission delay of the scheme is still very low and it is therefore suitable to use in delay-sensitive applications.

Paper D: Analog Distributed Source–Channel Coding Using Sinusoids [5]

Distributed source coding can be used to reduce transmission rate or, in the case of analog transmission, mitigate the effects of the channel noise. Similarly to Paper C, we look at side-information aware transmission based on analog mappings. In particular, we assume that an analog source is to be transmitted to a receiver that has access to correlated side information, as in the Wyner–Ziv problem. From the Cramér–Rao lower bound, we observe general properties of analog distributed source–channel mappings. It is especially clear how the stretch factor influences the performance and we use this insight to propose two different mappings based on sinusoidal waveforms. The proposed transmission scheme is numerically evaluated and shown to perform well, particularly in the low-SNR regime.

Paper E: Iterative Source–Channel Coding Approach to Witsenhausen's Counterexample [6]

In this paper, we make a detour to the area of automatic control and apply the methods we have used in joint source–channel coding to Witsenhausen's famous counterexample. In 1968, Witsenhausen showed that even in the simple linear quadratic static team decision problem, nonlinear decisions could outperform any given linear decision. This problem has served as a benchmark problem for decades where researchers try to achieve the optimal solution. The iterative source–channel coding approach we apply to this problem makes no assumptions of the shape of the mappings but still results in the lowest cost known to date.

Paper F: Optimized Low-Delay Source–Channel–Relay Mappings [7]

The three-node relay channel with a Gaussian source is studied for transmission subject to a low-delay constraint. We propose a joint source–channel coding design algorithm and evaluate the performance numerically. The designed systems are compared with reference systems, based on modular source and channel coding, and the distortion rate function for the Gaussian source using known achievable rates for the relay channel. The numerical comparisons show that the joint design works well and gives significantly better performance than the reference systems. By studying the structure of the optimized source–channel and relay mappings, we provide useful insights on how the optimized systems work. Interestingly, the design algorithm generally produces relay mappings with a structure that resembles Wyner–Ziv compression.

Paper G: Design and Performance of Optimized Relay Mappings [8]

Similarly to Paper F, we consider the three-node relay channel but now we study the transmission of a discrete information symbol. We let the relay be a memoryless function and formulate necessary conditions for the optimality of the relay mapping and the detector. Based on these, we propose a design algorithm to find relay mappings such that the symbol error rate at the destination is minimized. At virtually no extra complexity, the optimized relay mappings give remarkable power gains compared to the existing schemes detect-and-forward, amplify-and-forward, and estimate-and-forward. The proposed system is more flexible than all of the reference systems, since it finds a good tradeoff between soft and hard decisions depending on all link qualities. The optimized relay mappings are illustrated for different scenarios and the dependency between the relay mapping and the link qualities is discussed in detail.

Paper H: Closed-Form Sum-MSE Minimization for the Two-User Gaussian MIMO Broadcast Channel [9]

We consider sum-MSE minimization for the Gaussian multiple-input, multiple-output broadcast channel. Motivated by low-delay and low-complexity constraints, we turn to analog transmission using linear precoding, where the problem is to determine the optimal beamforming vector and power allocation to use for each user's message. By using recently discovered properties of this problem, we derive a closed-form expression for the optimal power allocation in the two-user scenario and propose a conceptually simple and efficient algorithm that handles an arbitrary number of users.

7 Conclusions and Future Work

In this thesis, methods for optimization of cooperative communication systems subject to very low delays have been considered. The principal contributions of this thesis can be summarized as follows

- Nonlinear is better than linear. Throughout the thesis, we show how nonlinear mappings outperform linear mappings for problems where there is side-information available. (Papers A–G)
- Practical optimized low-delay mappings for distributed source–channel coding are designed. (Papers A–D)
- A framework for optimization of otherwise difficult problems is developed. It is shown that it is possible to obtain good solutions by using relatively simple algorithms. (Papers A, B, E–G)
- Visualization of numerically optimized mappings. The visualization provides inspiration and insights into how to design structured low-delay systems. (Papers A, B, E–G)

The optimized systems that are produced in this thesis have good performance when compared to other systems of the same dimensionality, but the gap to OPTA is in many cases significant. Directions for future work include investigating how better bounds can be derived, that is, bounds that take the low-dimensionality of the system into account. The results in [90, 91] can potentially provide an opening in this direction. A tighter bound on OPTA for low-dimensional systems would strengthen the results in this thesis further.

References

[1] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, pp. 379–423, 623–656, July, October 1948.

[2] N. Wernersson, J. Karlsson, and M. Skoglund, "Distributed quantization over noisy channels," *IEEE Trans. on Communications*, vol. 57, no. 6, pp. 1693–1700, June 2009.

[3] J. Kron, F. Alajaji, and M. Skoglund, "Low-delay joint source–channel mappings for the Gaussian MAC," *IEEE Trans. on Communications*, submitted May 2011.

[4] J. Karlsson and M. Skoglund, "Lattice-based source–channel coding in wireless sensor networks," in *IEEE International Conference on Communications*, 2011, to appear.

[5] ——, "Analog distributed source–channel coding using sinusoids," in *International Symposium on Wireless Communication Systems*, September 2009, pp. 279–282, invited paper.

[6] J. Kron, A. Gattami, T. J. Oechtering, and M. Skoglund, "Iterative source–channel coding approach to Witsenhausen's counterexample," *Automatica*, submitted March 2011.

[7] J. Karlsson and M. Skoglund, "Optimized low-delay source–channel–relay mappings," *IEEE Trans. on Communications*, no. 5, pp. 1397–1404, May 2010.

[8] ——, "Design and performance of optimized relay mappings," *IEEE Trans. on Communications*, no. 9, pp. 2718–2724, September 2010.

[9] J. Kron, D. Persson, M. Skoglund, and E. G. Larsson, "Closed-form sum-MSE minimization for the two-user Gaussian MIMO broadcast channel," *IEEE Communications Letters*, submitted March 2011.

[10] J. Karlsson, N. Wernersson, and M. Skoglund, "Distributed scalar quantizers for noisy channels," in *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, April 2007, pp. 633–636.

[11] N. Wernersson, J. Karlsson, and M. Skoglund, "Distributed scalar quantizers for Gaussian channels," in *Proceedings IEEE Int. Symp. Information Theory*, June 2007, pp. 1741–1745.

[12] J. Karlsson and M. Skoglund, "Joint source–channel mappings for the relay channel," *Proc. of IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, pp. 2953–2956, April 2008.

[13] ——, "Design and performance of optimized relay mappings," in *5th International Conference on Broadband Communications, Networks and Systems*, September 2008.

[14] J. Karlsson, A. Gattami, T. J. Oechtering, and M. Skoglund, "Iterative source–channel coding approach to Witsenhausen's counterexample," in *Proceedings American Control Conference*, 2011, to appear.

[15] K. Kansanen, A. N. Kim, R. Thobaben, and J. Karlsson, "Low complexity bandwidth compression mappings for sensor networks," in *Proc. of Int. Symp. on Communications, Control and Signal Processing*, 2010.

[16] D. Persson, J. Karlsson, M. Skoglund, and E. G. Larsson, "Joint source-channel coding for the MIMO broadcast channel," *IEEE Trans. on Signal Processing*, submitted September 2010, revised February 2011.

[17] G. J. Pottie, "Wireless sensor networks," in *Information Theory Workshop, 1998*, June 1998, pp. 139–140.

[18] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks: a survey," *Computer networks*, vol. 38, no. 4, pp. 393–422, 2002.

[19] W. Dargie and C. Poellabauer, *Fundamentals of wireless sensor networks: theory and practice*. Wiley, 2010.

[20] L. Perez-Lombard, J. Ortiz, and C. Pout, "A review on buildings energy consumption information," *Energy and Buildings*, vol. 40, no. 3, pp. 394–398, 2008.

[21] *Energy efficiency in buildings: Facts and trends*, World Business Council for Sustainable Development, July 2008, available at www.wbcsd.org.

[22] *Smart home and intelligent building control: energy efficiency in buildings with ABB i-bus KNX*, ABB, 2009, order number 2CDC 500 060 M0201.

[23] V. C. Gungor and G. P. Hancke, "Industrial wireless sensor networks: Challenges, design principles, and technical approaches," *Industrial Electronics, IEEE Transactions on*, vol. 56, no. 10, pp. 4258–4265, October 2009.

[24] S. Kim, S. Pakzad, D. Culler, J. Demmel, G. Fenves, S. Glaser, and M. Turon, "Health monitoring of civil infrastructures using wireless sensor networks," in *Proceedings of the 6th international conference on Information processing in sensor networks*. ACM, 2007, pp. 254–263.

[25] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. Wiley-Interscience, 2006.

[26] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near shannon limit error-correcting coding and decoding: Turbo-codes," in *IEEE International Conference on Communications*, vol. 2, Geneva, Switzerland, May 1993, pp. 1064–1070.

[27] R. Gallager, "Low-density parity-check codes," *IRE Trans. on Information Theory*, vol. 8, no. 1, pp. 21–28, January 1962.

[28] D. J. C. MacKay, "Good error-correcting codes based on very sparse matrices," *IEEE Trans. on Information Theory*, vol. 45, no. 2, pp. 399–431, March 1999.

[29] D. Tse and P. Viswanath, *Fundamentals of wireless communication*. Cambridge Univ Pr, 2005.

[30] D. A. Huffman, "A method for the construction of minimum-redundancy codes," *Proceedings of the IRE*, vol. 40, no. 9, pp. 1098–1101, September 1952.

[31] J. Ziv and A. Lempel, "A universal algorithm for sequential data compression," *IEEE Trans. on Information Theory*, vol. 23, no. 3, pp. 337–343, May 1977.

[32] ——, "Compression of individual sequences via variable-rate coding," *IEEE Trans. on Information Theory*, vol. 24, no. 5, pp. 530–536, September 1978.

[33] H. L. V. Trees, *Detection, Estimation, and Modulation Theory. Part I.* Wiley, 1968.

[34] T. Berger, *Rate-distortion theory.* Englewood Cliffs, NJ, USA: Prentice Hall, 1971.

[35] S. P. Lloyd, "Least Squares Quantization in PCM," *Bell Telephone Laboratories Paper*, 1957.

[36] ——, "Least Squares Quantization in PCM," *IEEE Trans. on Information Theory*, vol. 28, no. 2, pp. 129–137, March 1982.

[37] J. Max, "Quantizing for minimum distortion," *IRE Trans. on Information Theory*, vol. 6, pp. 7–12, March 1960.

[38] Y. Linde, A. Buzo, and R. M. Gray, "An algorithm for vector quantizer design," *IEEE Trans. on Communications*, vol. 28, no. 1, pp. 84–95, January 1980.

[39] T. D. Lookabaugh and R. M. Gray, "High-resolution quantization theory and the vector quantizer advantage," *IEEE Trans. on Information Theory*, vol. 35, no. 5, pp. 1020–1033, September 1989.

[40] E. Agrell, T. Eriksson, A. Vardy, and K. Zeger, "Closest point search in lattices," *IEEE Trans. on Information Theory*, vol. 48, no. 8, pp. 2201–2214, August 2002.

[41] J. H. Conway and N. J. A. Sloane, *Sphere packings, lattices and groups.* New York, USA: Springer-Verlag, 1988.

[42] M. Skoglund, "On channel-constrained vector quantization and index assignment for discrete memoryless channels," *IEEE Trans. on Information Theory*, vol. 45, no. 7, pp. 2615–2622, November 1999.

[43] A. Mehes and K. Zeger, "Randomly chosen index assignments are asymptotically bad for uniform sources," *IEEE Trans. on Information Theory*, vol. 45, no. 2, pp. 788–794, March 1999.

[44] N. Rydbeck and C. W. Sundberg, "Analysis of digital errors in nonlinear PCM systems," *IEEE Trans. on Communications*, vol. 24, no. 1, pp. 59–65, January 1976.

[45] N. Farvardin and V. Vaishampayan, "Optimal quantizer design for noisy channels: An approach to combined source–channel coding," *IEEE Trans. on Information Theory*, vol. 33, no. 6, pp. 827–838, November 1987.

[46] H. Kumazawa, M. Kasahara, and T. Namekawa, "A construction of vector quantizers for noisy channels," *Electronics and Communications in Japan*, vol. 67, no. 4, pp. 39–47, January 1984.

[47] K. A. Zeger and A. Gersho, "Vector quantizer design for memoryless noisy channels," in *IEEE International Conference on Communications*, Philadelphia, USA, June 1988, pp. 1593–1597.

[48] N. Farvardin, "A study of vector quantization for noisy channels," *IEEE Trans. on Information Theory*, vol. 36, no. 4, pp. 799–809, July 1990.

[49] N. Farvardin and V. Vaishampayan, "On the performance and complexity of channel-optimized vectorquantizers," *IEEE Trans. on Information Theory*, vol. 37, no. 1, pp. 155–160, January 1991.

[50] M. Gastpar, B. Rimoldi, and M. Vetterli, "To code, or not to code: lossy source-channel communication revisited," *IEEE Trans. on Information Theory*, vol. 49, no. 5, pp. 1147–1158, May 2003.

[51] C. E. Shannon, "Communication in the presence of noise," *Proc. IRE*, pp. 10–21, January 1949.

[52] J. M. Wozencraft and I. M. Jacobs, *Principles of Communication Engineering*. Wiley, 1965.

[53] D. J. Sakrison, *Transmission of Waveforms and Digital Information*. Wiley, 1968.

[54] J. Ziv, "The behavior of analog commnunication systems," *IEEE Trans. on Information Theory*, vol. 16, no. 5, pp. 587–594, September 1970.

[55] V. Vaishampayan, "Combined source–channel coding for bandlimited waveform channels," Ph.D. dissertation, University of Maryland, 1989.

[56] A. Fuldseth and T. A. Ramstad, "Bandwidth compression for continuous amplitude channels based on vector approximation to a continuous subset of the source signal space," in *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Munich, Germany, April 1997, pp. 3093–3096.

[57] A. Fuldseth, "Robust subband video compression for noisy channels with multilevel signaling," Ph.D. dissertation, NTNU Trondheim, 1997.

[58] S. Chung, "On the construction of some capacity-approaching coding schemes," Ph.D. dissertation, MIT, 2000.

[59] T. A. Ramstad, "Shannon mappings for robust communication," *Telektronikk*, vol. 98, no. 1, pp. 114–128, 2002.

[60] F. Hekland, P. A. Floor, and T. A. Ramstad, "Shannon–Kotel'nikov mappings in joint source–channel coding," *IEEE Trans. on Communications*, vol. 57, no. 1, pp. 94–105, January 2009.

[61] E. Akyol, K. Rose, and T. Ramstad, "Optimal mappings for joint source channel coding," in *IEEE Information Theory Workshop (ITW)*, January 2010.

[62] Y. Hu, J. Garcia-Frias, and M. Lamarca, "Analog joint source channel coding using non-linear mappings and MMSE decoding," *IEEE Trans. on Communications*, to appear.

[63] X. Cai and J. W. Modestino, "Bandwidth expansion Shannon mapping for analog error-control coding," in *Information Sciences and Systems, 2006 40th Annual Conference on*, March 2006, pp. 1709–1712.

[64] P. A. Floor, T. A. Ramstad, and N. Wernersson, "Power constrained channel optimized vector quantizers used for bandwidth expansion," in *International Symposium on Wireless Communication Systems*, October 2007.

[65] B. Chen and G. W. Wornell, "Analog error-correcting codes based on chaotic dynamical systems," *IEEE Trans. on Communications*, vol. 46, no. 7, pp. 881–890, July 1998.

[66] V. Vaishampayan and S. I. R. Costa, "Curves on a sphere, shift-map dynamics, and error control for continuous alphabet sources," *IEEE Trans. on Information Theory*, vol. 49, no. 7, pp. 1658–1672, July 2003.

[67] N. Wernersson, M. Skoglund, and T. Ramstad, "Polynomial based analog source–channel codes," *IEEE Trans. on Communications*, vol. 57, no. 9, pp. 2600–2606, September 2009.

[68] U. Mittal and N. Phamdo, "Hybrid digital-analog (HDA) joint source-channel codes for broadcasting and robust communications," *IEEE Trans. on Information Theory*, vol. 48, no. 5, pp. 1082–1102, May 2002.

[69] M. Skoglund, N. Phamdo, and F. Alajaji, "Hybrid digital–analog source–channel coding for bandwidth compression/expansion," *IEEE Trans. on Information Theory*, vol. 52, no. 8, pp. 3757–3763, August 2006.

[70] D. Slepian and J. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. on Information Theory*, vol. 19, no. 4, pp. 471–480, July 1973.

[71] A. D. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. on Information Theory*, vol. 22, no. 1, pp. 1–10, January 1976.

[72] A. Wyner, "Recent results in the Shannon theory," *IEEE Trans. on Information Theory*, vol. 20, no. 1, pp. 2–10, January 1974.

[73] J. Garcia-Frias and Y. Zhao, "Compression of correlated binary sources using turbo codes," *IEEE Communications Letters*, vol. 5, no. 10, pp. 417–419, October 2001.

[74] A. D. Liveris, Z. Xiong, and C. N. Georghiades, "Compression of binary sources with side information at the decoder using LDPC codes," *IEEE Communications Letters*, vol. 6, no. 10, pp. 440–442, October 2002.

[75] S. S. Pradhan and K. Ramchandran, "Distributed source coding using syndromes (DISCUS): design and construction," *IEEE Trans. on Information Theory*, vol. 49, no. 3, pp. 626–643, March 2003.

[76] Z. Xiong, A. Liveris, S. Cheng, and Z. Liu, "Nested quantization and Slepian-Wolf coding: a Wyner-Ziv coding paradigm for i.i.d. sources," in *IEEE Workshop on Statistical Signal Processing*, September 2003, pp. 399–402.

[77] S. Pradhan and K. Ramchandran, "Generalized coset codes for distributed binning," *IEEE Trans. on Information Theory*, vol. 51, no. 10, pp. 3457–3474, October 2005.

[78] T. J. Flynn and R. M. Gray, "Encoding of correlated observations," *IEEE Trans. on Information Theory*, vol. 33, no. 6, pp. 773–787, November 1987.

[79] W. Lam and A. R. Reibman, "Design of quantizers for decentralized estimation systems," *IEEE Trans. on Communications*, vol. 41, no. 11, pp. 1602–1605, November 1993.

[80] D. Rebollo-Monedero, R. Zhang, and B. Girod, "Design of optimal quantizers for distributed source coding," in *Proceedings IEEE Data Compression Conference*, March 2003, pp. 13–22.

[81] A. Saxena and K. Rose, "Distributed predictive coding for spatio-temporally correlated sources," in *Proceedings IEEE Int. Symp. Information Theory*, June 2007, pp. 1506–1510.

[82] B. Liu and B. Chen, "Channel-optimized quantizers for decentralized detection in sensor networks," *IEEE Trans. on Information Theory*, vol. 52, no. 7, pp. 3349–3358, July 2006.

[83] A. Saxena, J. Nayak, and K. Rose, "On efficient quantizer design for robust distributed source coding," in *Proceedings IEEE Data Compression Conference*, March 2006, pp. 63–71.

[84] N. Wernersson and M. Skoglund, "Nonlinear coding and estimation for correlated data in wireless sensor networks," *IEEE Trans. on Communications*, vol. 57, no. 10, pp. 2932–2939, October 2009.

[85] E. C. van der Meulen, "Three-terminal communication channels," *Adv. Appl. Prob.*, vol. 3, no. 1, pp. 120–154, 1971.

[86] M. A. Khojastepour, A. Sabharwal, and B. Aazhang, "On capacity of Gaussian 'cheap' relay channel," in *Proceedings IEEE Global Telecommunications Conference*, vol. 3, December 2003, pp. 1776–1780.

[87] T. Cover and A. E. Gamal, "Capacity theorems for the relay channel," *IEEE Trans. on Information Theory*, vol. 25, no. 5, pp. 572–584, September 1979.

[88] A. Høst-Madsen and J. Zhang, "Capacity bounds and power allocation for wireless relay channels," *IEEE Trans. on Information Theory*, vol. 51, no. 6, pp. 2020–2040, June 2005.

[89] A. E. Gamal, M. Mohseni, and S. Zahedi, "Bounds on capacity and minimum energy-per-bit for AWGN relay channels," *IEEE Trans. on Information Theory*, vol. 52, no. 4, pp. 1545–1561, April 2006.

[90] J. Ziv and M. Zakai, "On functionals satisfying a data-processing theorem," *IEEE Trans. on Information Theory*, vol. 19, no. 3, pp. 275–283, May 1973.

[91] A. Ingber, I. Leibowitz, R. Zamir, and M. Feder, "Distortion lower bounds for finite dimensional joint source-channel coding," in *Proceedings IEEE Int. Symp. Information Theory*, July 2008, pp. 1183–1187.

Part II

Included Papers

Paper A

Distributed Quantization over Noisy Channels

Niklas Wernersson, Johannes Karlsson, and Mikael Skoglund

Published in *IEEE Transactions on Communications,* June 2009.

Distributed Quantization over Noisy Channels

Niklas Wernersson, Johannes Karlsson, and Mikael Skoglund

Abstract

The problem of designing simple and energy-efficient sensor nodes in a wireless sensor network is considered from a joint source–channel coding perspective. An algorithm for designing distributed scalar quantizers for orthogonal channels is proposed and evaluated. In particular the cases of the binary symmetric channel as well as the additive white Gaussian noise channel are studied. It is demonstrated that correlation between sources can be useful in order to reduce quantization distortion as well as protecting data when being transmitted over non-ideal channels. It is also demonstrated that the obtained system is robust against channel SNR mismatch.

Index Terms–Source coding, quantization, channel coding, correlation.

1 Introduction

Wireless sensor networks are expected to play an important role in tomorrow's sensing systems. One important property in these networks is that there may be a high correlation between different sensor measurements due to high spatial density of sensor nodes. This motivates source coding of correlated sources, which has been analyzed in for instance [1] where the well known Slepian–Wolf theorem is stated. Ideas on how to perform practical Slepian–Wolf coding are presented in [2,3], allowing the use of powerful channel codes such as LDPC and Turbo codes in the context of distributed source coding, see e.g. [4,5]. For the case with continuous sources, i.e. lossy coding, relevant references include [6, 7]. In general, these methods require the use of long codes and the encoding complexity will require some data processing in the sensor nodes. This will therefore counteract one of the desired design criteria in sensor network design, namely low cost and energy efficient sensor nodes. In addition, in many applications for example in networked control, a low delay is essential, preventing the use of long codes.

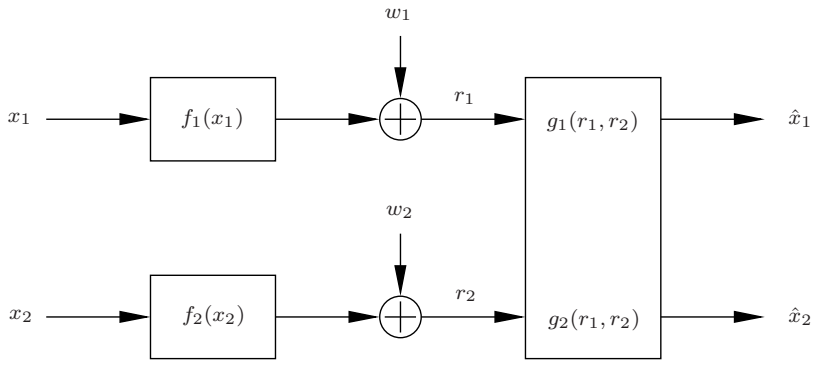


Figure 1: Structure of the system.

An alternative is therefore to design sensor nodes of very low complexity and low delay. This can be accomplished by interpreting the distributed source coding problem as a quantization problem. Previously, quantization of correlated sources has been studied in [8–13]. Our work is however targeted towards wireless sensor networks and introducing noisy channels is necessary in order to make the system more realistic. For non-ideal channels related previous work includes [14] which considers the problem of distributed detection over non-ideal channels. In [15] quantization of correlated sources in a packet network is studied, resulting in a general problem including multiple description coding as well as distributed source coding as special cases.

We will in this paper summarize and continue the work carried out in [16, 17] where distributed scalar quantizers were designed for different channel models. In what follows, we propose a design algorithm that results in sensor nodes operating on a sample by sample basis in a similar fashion as a channel optimized scalar quantizer (COSQ) [18].

2 Problem Formulation

We consider the problem of distributed joint source–channel coding illustrated in Figure 1. Two correlated random variables X_1 and X_2 are to be encoded by two encoders separated in space preventing cooperation between the encoders. To achieve low-complexity and low-delay encoding, the mappings f_1 and f_2 work in the following manner: f_1 and f_2 will first scalar quantize X_1 and X_2 to indexes i_1 and i_2 according to

$$q_k : \mathcal{X}_k \to \mathcal{I}_k \in \{0, 1, ..., N-1\} \quad \forall k \in \{1, 2\}. \tag{1}$$

and these indexes are then transmitted over an additive white Gaussian noise (AWGN) channel. Two different transmission methods will be studied resulting in two different channel models. The first model is created by using BPSK on the AWGN channel in conjunction with hard decision decoding. This results in a binary symmetric channel (BSC) with some given bit error probability. Hence, in the first model each index is transmitted using the BSC $R \triangleq \log_2 N$ times. In the second model we will transmit each index by mapping the quantization index to a symbol in an N pulse amplitude modulated (N–PAM) signal. We will refer to this case as the 'Gaussian channel'. We explain these two cases in greater detail below.

2.1 Binary Symmetric Channel

For the case of the BSC the quantization index i_k from (1) will be mapped to its binary representation as

$$\mathbf{f}_k : \mathcal{X}_k \stackrel{q_k}{\to} \mathcal{I}_k \to \{-1, 1\}^R \quad \forall k \in \{1, 2\}. \tag{2}$$

Hence, $\mathbf{f}_k$ will use q_k to create the index i_k which is then represented binary. These bits are transmitted over a Gaussian channel using BPSK resulting in

$$\mathbf{r}_k = \mathbf{f}_k(x_k) + \mathbf{w}_k \quad \forall k \in \{1, 2\} \tag{3}$$

where $\mathbf{w}$ is zero mean i.i.d. Gaussian noise with covariance matrix $\sigma_w^2\mathbf{I}$. For each of these R received values a hard decision decoding rule is applied such that

$$j_k(m) = \text{sign}(r_k(m)) \quad m = 1, 2, \cdots, R \tag{4}$$

where

$$\text{sign}(x) = \begin{cases} 1, & x \geq 0 \\ -1, & x < 0. \end{cases} \tag{5}$$

Given that -1 was transmitted, and letting $Q(\cdot)$ denote the Q-function, this will result in a bit error probability

$$\epsilon = \int_0^\infty \frac{1}{\sqrt{2\pi\sigma_w^2}} e^{-\frac{(r+1)^2}{2\sigma_w^2}} \mathrm{d}r = Q\left(\frac{1}{\sigma_w}\right), \tag{6}$$

which is also, due to the symmetry, the total bit error probability.

Denoting the decimal representation of $\mathbf{j}_k$ as j_k the decoding will be performed as

$$\hat{x}_k = g_k(j_1, j_2) \quad \forall k \in \{1, 2\}. \tag{7}$$

Hence, the decoding is based on both j_1 and j_2.

Given this system, we define the mean squared error (MSE) as

$$D = \frac{1}{2}(D_1 + D_2) = \frac{1}{2}\left(\mathbf{E}[(X_1 - \hat{X}_1)^2] + \mathbf{E}[(X_2 - \hat{X}_2)^2]\right) \tag{8}$$

and our objective is to design the encoders and the decoder in order to minimize the MSE.[1]

2.2 Gaussian Channel

For the Gaussian channel each of the indexes (i_1, i_2) are mapped to an N pulse amplitude modulated (N–PAM) signal such that

$$f_k(x_k) = \alpha(2q_k(x_k) - N + 1) \quad \forall k \in \{1, 2\}. \tag{9}$$

Here α is a constant such that the power constraints

$$\mathbf{E}[f_k(X_k)^2] \leq P \quad \forall k \in \{1, 2\} \tag{10}$$

are satisfied. The two PAM signals are then transmitted over two orthogonal channels, created by using e.g. TDMA or FDMA, resulting in the received values

$$r_k = f_k(x_k) + w_k \quad \forall k \in \{1, 2\} \tag{11}$$

where the noise terms w_k are independent zero-mean Gaussian distributed with variance σ_w^2. The decoder will have access to both r_1 and r_2 and forms its estimate of the original source data as

$$\hat{x}_k = g_k(r_1, r_2) \quad \forall k \in \{1, 2\}. \tag{12}$$

Here the objective is to design the encoders and the decoder in order to minimize the (MSE) from (8) under the power constraints given in (10).

3 Analysis

As in traditional Lloyd-Max training [19] we will optimize each part in the system in an iterative fashion keeping the other parts fixed. Note that the system contains three parts: two encoders and one decoder, although the decoder contains two decoding functions. We will in this section consider the design of these parts under the assumption that

$$X_k = Y + Z_k \quad \forall k \in \{1, 2\} \tag{13}$$

where Y, Z_1 and Z_2 are independent zero-mean Gaussian distributed random variables with variances σ_Y^2, $\sigma_{Z_1}^2 = \sigma_{Z_2}^2 = \sigma_Z^2$. Hence, X_1 and X_2 are correlated which can be exploited in the encoding as well as the decoding.

[1] As pointed out by one of the reviewers one could also define a weighted MSE as $D(\rho) = \rho D_1 + (1-\rho)D_2$ and adopt our derived equations accordingly. One interesting case would be $D(1)$ meaning that the second observation x_2 only serves as side information when estimating x_1. However, we will only study the case $D(0.5)$, i.e. as in (8).

For this jointly Gaussian distribution we get the conditional pdf

$$p(x_2|x_1) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\left(x_2 - \frac{\sigma_Y^2}{\sigma_Y^2+\sigma_Z^2} x_1\right)^2}{2\sigma^2} \right) \tag{14}$$

where

$$\sigma^2 = \frac{\sigma_Z^4 + 2\sigma_Y^2\sigma_Z^2}{\sigma_Y^2 + \sigma_Z^2}. \tag{15}$$

Without loss of generality we will further assume that $\mathbf{E}[X_1^2] = \mathbf{E}[X_2^2] = 1$, hence $\sigma_Y^2 + \sigma_Z^2 = 1$.

3.1 Encoder for BSC

Only the design of $\mathbf{f}_1$ will be considered since $\mathbf{f}_2$ can be designed in the same fashion. Given that the encoder $\mathbf{f}_1$ observes x_1 and produces index i_1 it can derive the expected distortions for D_1 and D_2 as

$$D_1(x_1, i_1) = \sum_{j_1} \sum_{j_2} P(j_1|i_1) P(j_2|x_1) \left[x_1 - g_1(j_1, j_2)\right]^2 \tag{16}$$

$$D_2(x_1, i_1) = \int \sum_{j_1} \sum_{j_2} P(j_1|i_1) p(x_2|x_1) P(j_2|q_2(x_2)) \left[x_2 - g_2(j_1, j_2)\right]^2 \mathrm{d}x_2, \tag{17}$$

where the integral is taken from $-\infty$ to ∞ and

$$P(j_2|x_1) = \sum_{i_2} P(j_2|i_2) P(i_2|x_1) \tag{18}$$

where

$$P(i_2|x_1) = \int\limits_{x_2 : q_2(x_2) = i_2} p(x_2|x_1) \mathrm{d}x_2. \tag{19}$$

The other transition probabilities $P(\cdot|\cdot)$ are straightforward to derive, see e.g. [20]. In order to minimize the distortion (8) the quantizer $q_1(x_1)$ should be designed according to

$$q_1(x_1) = \arg\min_{i_1} (D_1(x_1, i_1) + D_2(x_1, i_1)). \tag{20}$$

In [18] the case of a single source was studied. In this case, the solution resulted in encoder regions which were intervals, and analytical expressions for finding the endpoints of these intervals were derived. However, (20) does in general not result in a similar solution and the encoder regions will in general *not be intervals*, but rather unions of separated intervals (this will be illustrated in Section 4).

3.2 Encoder for Gaussian Channel

For the Gaussian channel D_1 and D_2 can be expressed as

$$D_1(x_1, i_1) = \iint p(r_1|i_1)p(r_2|x_1)\left[x_1 - g_1(r_1, r_2)\right]^2 \mathrm{d}r_2\mathrm{d}r_1 \tag{21}$$

$$D_2(x_1, i_1) = \iiint p(r_1|i_1)p(x_2|x_1)p(r_2|q_2(x_2))\left[x_2 - g_2(r_1, r_2)\right]^2 \mathrm{d}r_2\mathrm{d}x_2\mathrm{d}r_1, \tag{22}$$

where the integrals are taken from $-\infty$ to ∞. In order to minimize the distortion (8) under the power constraint (10) the quantizer $q_1(x_1)$ should be designed according to

$$q_1(x_1) = \arg\min_{i_1}(D_1(x_1, i_1) + D_2(x_1, i_1) + \lambda(2i_1 - N + 1)^2). \tag{23}$$

Here, the first two terms aim at minimizing the distortion introduced by the quantizer whereas the third term will allow us to control the power consumption by choosing a value for the Lagrangian multiplier λ, see e.g. [21]. Unfortunately the integrals in (21)–(22) are difficult to evaluate since they contain $g_1(r_1, r_2)$ and $g_2(r_1, r_2)$ which vary with r_1 and r_2. In order to get around this problem we use the technique of prequantizing r_1 and r_2 according to

$$h : (\mathcal{R}_1, \mathcal{R}_2) \rightarrow (\mathcal{J}_1, \mathcal{J}_2) \in \{1, 2, ..., M\}^2 \tag{24}$$

which will produce the decoding functions

$$\hat{x}_k = g_k(h(r_1, r_2)) = g_k(j_1, j_2) \quad \forall k \in \{1, 2\}. \tag{25}$$

Furthermore, in this work we choose $M = N$ and let $h(r_1, r_2)$ simply map (r_1, r_2) to the closest possible output from the encoders defined by $(f_1(x_1), f_2(x_2))$. Hence

$$h(r_1, r_2) = \arg\min_{(j_1, j_2)}((r_1 - \alpha(2j_1 - N + 1))^2 + (r_2 - \alpha(2j_2 - N + 1))^2). \tag{26}$$

The decoding functions will now be piecewise linear over r_1 and r_2 which greatly simplifies the derivation of (21–22) and we get the same equations as in (16–17) (although the transition probabilities are different). Using (23) together with (16–17) will therefore define the optimal quantizer $q_1(x_1)$ under the assumption that the decoder and the second encoder are fixed.

3.3 Decoder

Assuming fixed encoders it is a well known fact from estimation theory that the optimal, in minimum MSE sense, estimates of x_1 and x_2 are given as

$$\hat{x}_k = g_k(j_1, j_2) = \mathbf{E}[x_k \,|\, j_1, j_2] \quad \forall k \in \{1, 2\}. \tag{27}$$

Hence, (27) is used to derive the decoders for both considered transmission methods.

3.4 Design algorithm

Based on the developed equations (20), (23) and (27) it will be possible to optimize the encoders and the decoder. A natural order to optimize these is: *1)* the first encoder, *2)* the decoder, *3)* the second encoder, *4)* the decoder. Each step in the iteration will guarantee the distortion to decrease and the training is repeated until the solution converges. Just as in the case of the Lloyd-Max algorithm this will result in a locally optimal system which is not necessarily the global optimum.

One problem with the suggested training above is that the obtained local optimum produced will depend greatly on the initialization of the decoder and encoders. In fact, in our simulations we experienced that very poor local optima were often found using the approach suggested above. This problem has also been encountered in [18, 22–24] where the method of noisy channel relaxation was introduced. The idea is essentially that it is easier to find a good local optimum for channels with high noise energy than for channels with low noise energy. Therefore a system is first designed for a very bad channel. Next, the channel quality is gradually improved and a new system is designed in each step. For each design, a full iterative training algorithm is executed using the reconstruction codebook from the previous design as initialization for the current design. We incorporate this idea by starting designing a system for a noise variance $\sigma_w'^2 \gg \sigma_w^2$. When this is completed $\sigma_w'^2$ is decreased with a stepsize σ_Δ^2 and a new system is designed. This is repeated L times. The algorithm is summarized below.

1. Initialize encoders and optimize the decoder by using (27).
2. Set values for L and σ_Δ^2. Create $\sigma_w'^2 = \sigma_w^2 + L\sigma_\Delta^2$.
3. Design a system for the channel noise $\sigma_w'^2$ according to:
   (a) Set the iteration index $k = 0$ and $D^{(0)} = \infty$.
   (b) Set $k = k + 1$.
   (c) Find the optimal quantizer q_1 by using (20) (or (23)).
   (d) Find the optimal decoder by using (27).
   (e) Find the optimal quantizer q_2 by using $q_2(x_2)$'s equivalence to (20) (or (23)).
   (f) Find the optimal decoder by using (27).
   (g) Evaluate the distortion $D^{(k)}$ for the system. If the relative improvement of $D^{(k)}$ compared to $D^{(k-1)}$ is less than some threshold $\delta > 0$ go to Step 4. Otherwise go to Step (b).

4. If $\sigma_w'^2 = \sigma_w^2$ stop the iteration. Otherwise create $\sigma_w'^2 = \sigma_w'^2 - \sigma_\Delta^2$ and go to Step 3 using the current encoders and decoder when initializing the next iteration.

We also experienced that when searching for a good local optima a small improvement was sometimes obtained by also performing a noise relaxation procedure for the correlation, i.e. varying σ_z^2. However, the main improvement was obtained by the algorithm above.

3.5 Optimal Performance Theoretically Attainable

Recently the rate region for the quadratic two-terminal source coding problem has been completely characterized in [25]. Furthermore, in [26] it is shown that separating the source and channel code design, when the block lengths are approaching infinity, will be asymptotically optimal for the problem we are considering. Hence, by simply studying the channel capacity of the different orthogonal channels we get rate constraints, R_1 and R_2, on the source code since these rates can be safely communicated to the decoder. Assuming that we have access to a capacity achieving channel code for the BSC we can transmit

$$\beta_{\text{BSC}} = R_1 = R_2 = RC_{\text{BSC}} = R(1 + \epsilon \log_2 \epsilon + (1-\epsilon)\log_2(1-\epsilon)) \quad (28)$$

on each channel, here C_{BSC} is the capacity of the BSC [27]. For the Gaussian channel we note that both encoders have the same power constraint (10) and that both channels have the same noise power. This gives

$$\beta_{\text{AWGN}} = R_1 = R_2 = C_{\text{AWGN}} = \frac{1}{2}\log_2\left(1 + \frac{P}{\sigma_w^2}\right) \quad (29)$$

where C_{AWGN} is the capacity of the AWGN channel [27]. Using the appropriate β, from (28) or (29), and simplifying the expressions in [25] (remember the assumption $\sigma_Y^2 + \sigma_Z^2 = 1$) gives

$$D_1 D_2 \geq 2^{-4\beta}(1 - \sigma_Y^4) + \sigma_Y^4 2^{-8\beta}. \quad (30)$$

Since D_1 is inversely proportional to D_2 the total distortion in (8) will be minimized by setting $D = D_1 = D_2$. This gives the optimal performance theoretically attainable (OPTA) according to

$$D = \sqrt{2^{-4\beta}(1 - \sigma_Y^4) + \sigma_Y^4 2^{-8\beta}}. \quad (31)$$

That is, D in (31) is the lowest possible achievable distortion for this problem.

4 Simulations

We will here visualize the structure of the encoders obtained when using the design algorithm presented in Section 3.4. The performance of a designed system is also compared to the OPTA derived in Section 3.5. In order to do so we measure the signal-to-distortion ratio (SDR) defined as

$$\mathrm{SDR} = 10\log_{10}\left(\frac{\mathbf{E}[X_1^2]+\mathbf{E}[X_2^2]}{\mathbf{E}[(X_1-\hat{X}_1)^2]+\mathbf{E}[(X_2-\hat{X}_2)^2]}\right) \tag{32}$$

and we also define the correlation SNR as

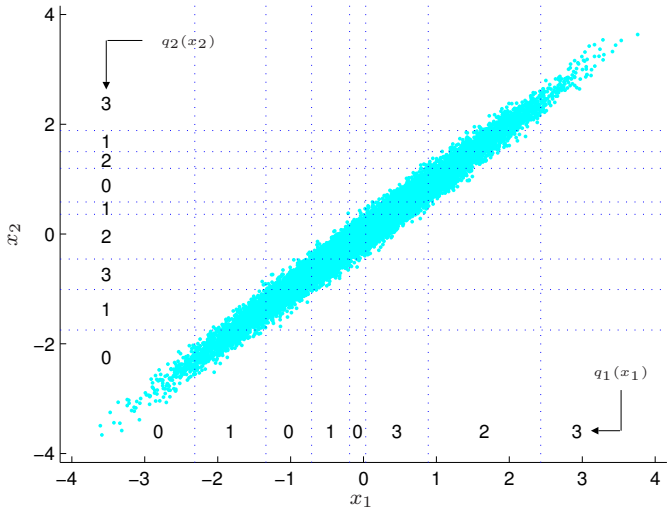
$$\mathrm{CSNR} = 10\log_{10}\left(\frac{\sigma_Y^2}{\sigma_Z^2}\right). \tag{33}$$

Hence, CSNR $= -\infty$ dB means that X_1 and X_2 are uncorrelated and CSNR $= \infty$ dB means that they are fully correlated. We use the term SNR when referring to the channel SNR defined as $10\log_{10}(P/\sigma_w^2)$. As initial encoders we used uniform quantizers and for the case of BSC the folded binary code [28] was used as initial codeword assignment.

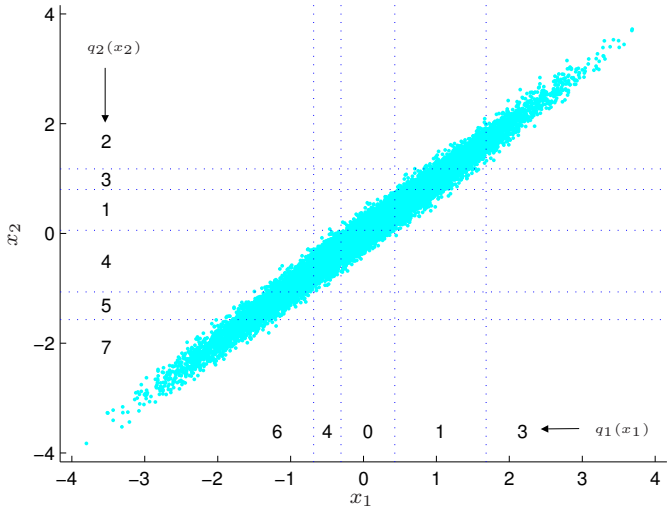
4.1 Structure of the Codebook - BSC

In Figure 2 systems have been designed for CSNR $= 20$ dB and the resulting encoders are illustrated for different bit error probabilities. Starting with Figure 2(a), where $\epsilon = 0$ and $R = 2$ bits per sample and source, a number of source data samples are marked by the grayish distribution. These samples are spread out along the diagonal due to the correlation between x_1 and x_2. In the plot the different quantization intervals for q_1 and q_2 are marked by the dashed lines. The representation of codewords produced by the quantizers in the different intervals are also marked. It is here interesting to note that many of the codewords are used for more than one quantization region. For example the codeword $i_2 = 1$ is used for 3 separated intervals such that $q_2(x_2) = 1$ when x_2 belongs (approximately) to the set $\{(-1.7,-1.0)\cup(0.4,0.6)\cup(1.5,1.9)\}$. With information from only one of the channels it is not possible to identify which of these different intervals x_2 belongs to. However, with help from i_1 (or rather j_1) this can be accomplished since $i_1 = 0$ or 1 is highly likely when $x_2 \in (-1.7,-1.0)$, $i_1 = 3$ is highly likely when $x_2 \in (0.4,0.6)$, and so on. Hence, i_1 will indicate which of the separated intervals x_2 belongs to. In this way the distributed coding is used to decrease the quantization distortion. It is noteworthy that the sets of separated intervals are created by the design algorithm despite the fact that the initial encoders are regular quantizers where all quantization regions are single intervals.

When the bit error probability increases the encoders will be more restrictive in using all possible codewords since they will be more likely to be



(a)



(b)

Figure 2: Encoder structures for systems with CSNR = 20 dB and R = 2 bits/sample, $\epsilon = 0$ in (a) and R = 3 bits/sample, $\epsilon = 0.05$ in (b). The small dots in the background show a sample distribution of (X_1, X_2) and the dashed lines show the boundaries for the quantization regions.

decoded incorrectly. In Figure 2(b) a system has been designed for $\epsilon = 0.05$ and $R = 3$ bits per sample and source. As can be seen only a subset of the codewords are now used by the encoders and these codewords have been placed with an appropriate index assignment.

4.2 Structure of the Codebook - Gaussian Channel

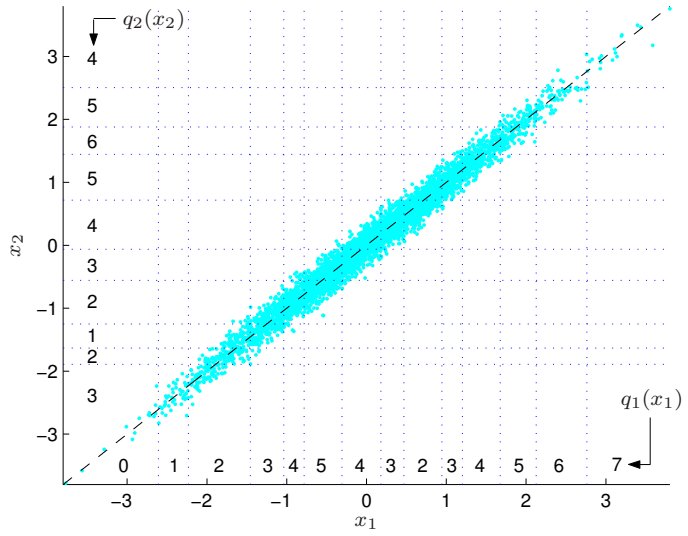
In order to illustrate the characteristics of the resulting system for the Gaussian channel a simple system with $N = 8$ has been designed and used for SNR = 10 dB and CSNR = 20 dB. The resulting quantizers are shown in Figure 3(a) and in Figure 3(b) it is illustrated how the quantization indexes are mapped to the channel space.

Starting with Figure 3(a) we once again see that the codewords will be reused as discussed in the previous section. See for instance the codeword $i_1 = 5$ which is used both when x_1 belongs (approximately) to the set $(-0.8, -0.3) \cup (1.7, 2.1)$. With help from i_2 (or rather r_2) the decoder will be able to distinguish between these two intervals since $i_2 = 2$ or 3 is highly likely if x_1 belongs to the first interval and otherwise $i_2 = 5$ or 6 will be highly likely.

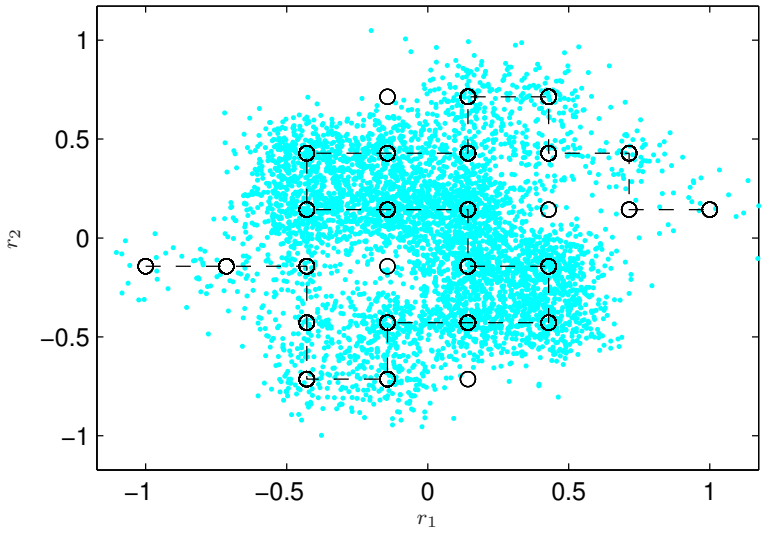
Let us now consider what will happen when the source data is quantized by q_1 and q_2 and mapped to the signal space by f_1 and f_2 as described by (9). Both f_1 and f_2 uses 8–PAM, resulting in 64 possible combinations at the encoder output. However, many of these combinations are very unlikely to occur and for the simulation conducted the occurred outputs are marked by circles in Figure 3(b). Furthermore, when transmitting these output values the channels will add noise to the outputs creating a distribution of (r_1, r_2) which is indicated by the grayish distribution in Figure 3(b).

Finally, some extra source data values were created in Figure 3(a) where $x_1 = x_2 = x$, hence $\sigma_Z^2 = 0$, and we let x increase from $-\infty$ to ∞. These values are marked by the line along the diagonal. The reason for adding these extra fully correlated values is that studying how this line is mapped to the channel signal space will give insight in the mapping procedure. By connecting the outputs created, when encoding this extra source data, we see how the line is mapped to the channel signal space (marked by a dashed line in Figure 3(b)). From this we note that, in general, samples far apart in the source signal space are also far apart in the channel signal space and vice versa. The power constraint will also focus the outputs in the area around the origin as much as possible in order to keep down the power consumption.

In Figures 4(a)–4(b) we present two other illustrations of the channel space, this time for $N = 32$. Figure 4(a) represents the case of high CSNR whereas Figure 4(b) represents low CSNR. From, for instance, Figure 4(a) we can imagine an underlying continuous curve $(f_1'(x), f_2'(x))$ which would be a good choice if we let $N \to \infty$. Furthermore, the curves created by

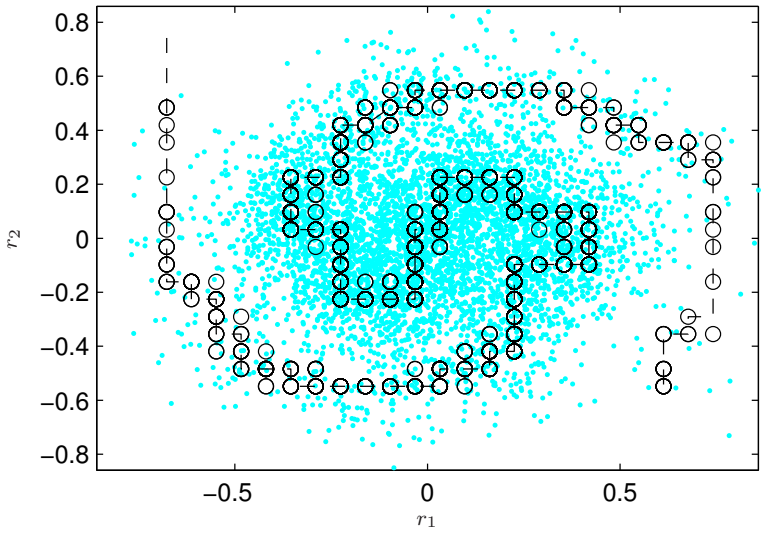


(a)

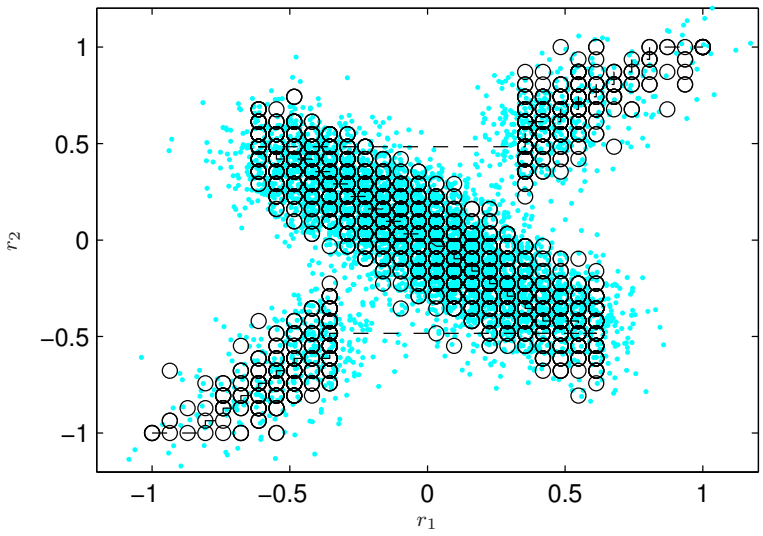


(b)

Figure 3: (a) Quantizers and (b) the corresponding mapping to the channel space for a system designed and used for $N = 8$, SNR $= 10$ dB and CSNR $= 20$ dB.



(a)



(b)

Figure 4: Other mappings to the channel space are illustrated for (a) $N = 32$, SNR = 7.2 dB and CSNR = 30 dB and (b) $N = 32$, SNR = 10 dB and CSNR = 13 dB.

$(f_1(x), f_2(x))$ appear to, especially for the high CSNR case, relate to what is often referred to as the bandwidth expansion problem mentioned already in one of Shannon's first papers [29]. This is the resulting problem when CSNR $= \infty$ dB, i.e. $x_1 = x_2$, meaning that one is allowed to use a channel twice in order to transmit one source sample. It is well know that optimal encoding functions f_1 and f_2 will be nonlinear for this case, see e.g. [30, 31] and the references therein.

The connection between the bandwidth expansion problem and distributed source coding is an interesting insight and we draw the conclusion that if an analog system is to be used for distributed source coding linear operations for f_1 and f_2 are not necessarily appropriate. We have elaborated on this further in [32]. It is interesting to note that the curves $(f_1(x), f_2(x))$ are not necessarily continuous when $N \to \infty$ which also seems to be indicated by Figure 4(b).

Finally we comment on the fact that the number of used encoder outputs from f_1 and f_2 are not the same. For instance, in Figure 3(b) f_1 uses 8 encoder outputs whereas f_2 only uses 6. The curves $(f_1(x), f_2(x))$ created will have two properties, the first is that the distance between different folds of the curve will be high enough to combat the channel noise. The second property is that the created curves will place the most commonly occurring encoder outputs in the center where the power consumption is low. Less common encoder outputs will be placed further out and the curves will therefore grow outwards. However, due to the power constraint the power consumption will at some stage become to high and the algorithm will prevent the curve to grow any further. This will therefore cause the encoders to use different numbers of outputs.

4.3 Performance Evaluation

We begin with evaluating a system designed for the BSC with $R = 3$ bits per source sample, $\epsilon = 0.01$ which is equivalent to a channel with SNR $= 7.3$ dB (using the inverse of (6)) and CSNR $= 13$ dB. In Figure 5 we study the performance of the system (dashed line) when the SNR is varied. We have also included the OPTA (solid line) as well as a reference method (dotted line) in the plot. The reference method is traditional COSQ [18] where two independent COSQ's are designed for $R = 3$ bits per sample and SNR $= 7.3$ dB, hence the correlation is not taken into consideration in the design. At the design SNR the gap to the OPTA curve is about 8 dB. Here it should be emphasized that achieving the OPTA requires infinite block-lengths, while our system works without delay on a sample by sample basis. Also, achieving OPTA will require that the system is optimized for each specific SNR whereas our simulated system is designed for one particular SNR but used for all simulated SNR's. By comparing to the reference method we can see that the gain of utilizing the source correlation in the

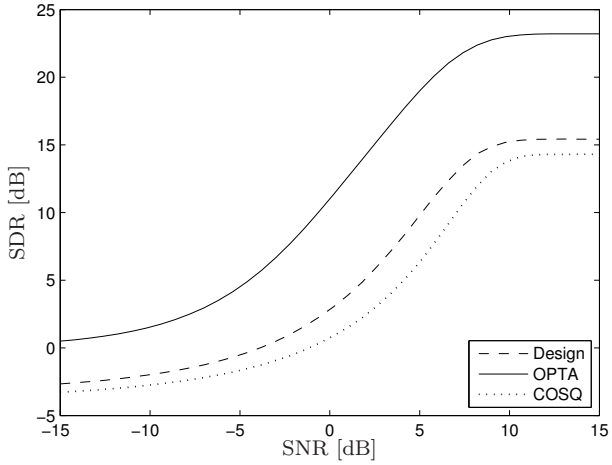


Figure 5: Evaluating the effect of varying the SNR when CSNR = 13 dB for a system designed for $R = 3$ bits per sample, SNR = 7.3 dB and CSNR = 13 dB.

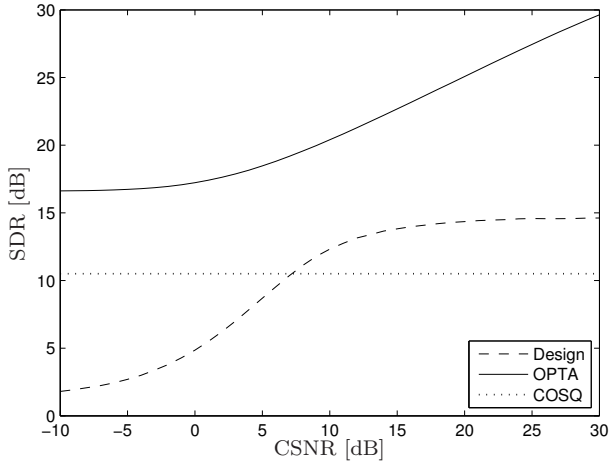


Figure 6: Evaluating the effect of varying the CSNR when SNR = 7.3 dB for a system designed for $R = 3$ bits per sample, SNR = 7.3 dB and CSNR = 13 dB.

encoders and the decoder is about 3 dB at the design SNR. When the SNR is increased above 10 dB the main contribution to the distortion comes from quantization which is limited by $R = 3$ bits per source sample, increasing the SNR above this point will therefore only have a small influence on the performance.

Next we keep the SNR fixed at 7.3 dB and look at the effect of a CSNR mismatch. That is, we evaluate the performance of the same system as above, which is designed for CSNR = 13 dB, when the true source correlation is varied. The result is shown in Figure 6 where we can see that the system is quite sensitive to a too low CSNR whereas a higher CSNR only gives a slight improvement in the performance. The designed system is however better than the reference method as long as the CSNR is above 7 dB. The reference method will not depend on the correlation and therefore has a constant performance.

In Figures 7–8 we present similar simulation results for the Gaussian channel. The simulated system is the same system as shown in Figure 4(b) designed for $N = 32$, SNR = 10 dB, CSNR = 13 dB and $\lambda = 0.01$. We have also here included the OPTA as well as a reference method, traditional COSQ, in the plot. In Figure 7 we let the CSNR equal 13 dB, hence what the system is designed for, but we vary the true SNR in order to study the effects of SNR mismatch. From the figure we see that in the area around SNR = 10 dB we are about 4 dB away from the OPTA (the additional figure is a magnification of the region around SNR = 10 dB). Increasing the SNR from this point will naturally increase the performance of OPTA and lowering the SNR will decrease the performance. It is therefore interesting to note that the designed system is able to follow the OPTA curve with essentially a constant 4 dB distance in the interval SNR $\in$ [5 dB, 15 dB]. The system is hence robust to a too low SNR and at the same time it is able to exploit a high SNR in order to increase the performance. Comparing the system to the reference method we see that there is about a 1 dB performance gain when the SNR is above 5 dB.

In Figure 8 we instead let SNR=10 dB and study the effect of a mismatch in CSNR. Here it appears as the system is, just as in the BSC case, more sensitive to a too low CSNR. It can tolerate some mismatch but the performance will quite soon start decreasing rapidly. A too high CSNR only gives a slight improvement in performance and a saturation level is reached after only a few dB increase. Hence, for a high CSNR the proposed method has the better performance and vice versa.

5 Conclusions

A design algorithm for joint source–channel optimized distributed scalar quantizers is presented and evaluated. The resulting system works on a

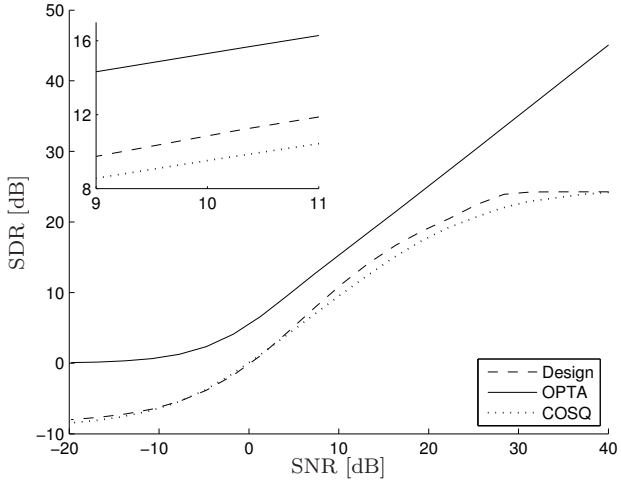


Figure 7: Evaluating the effect of varying the SNR when CSNR = 13 dB for a system designed for SNR = 10 dB and CSNR = 13 dB. The upper left plot shows a magnification of the area around SNR = 10 dB.

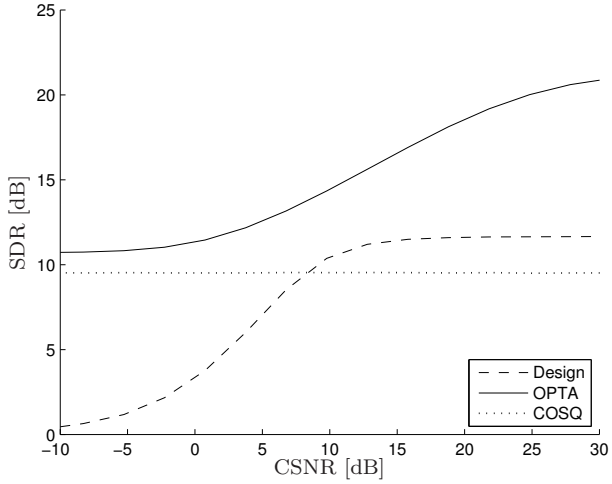


Figure 8: Evaluating the effect of varying the CSNR when SNR = 10 dB for a system designed for SNR = 10 dB and CSNR = 13 dB.

sample by sample basis yielding a very low encoding complexity, at an insignificant delay. Due to the source correlation, the resulting quantizers use the same codeword for several separated intervals in order to reduce the quantization distortion. Furthermore, the resulting quantization indexes are mapped to the channel signal space in such a way that source samples far from each other in the source signal space are well separated also in the channel signal space, and vice versa. This gives systems robust against channel SNR mismatch which was shown when comparing designed systems to the optimal performance theoretically attainable. The proposed main application of these quantizers is in low-complexity and energy-efficient wireless sensor nodes.

References

[1] D. Slepian and J. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. on Information Theory*, vol. 19, no. 4, pp. 471–480, July 1973.

[2] A. Wyner, "Recent results in the Shannon theory," *IEEE Trans. on Information Theory*, vol. 20, no. 1, pp. 2–10, January 1974.

[3] S. S. Pradhan and K. Ramchandran, "Distributed source coding using syndromes (DISCUS): design and construction," *IEEE Trans. on Information Theory*, vol. 49, no. 3, pp. 626–643, March 2003.

[4] A. D. Liveris, Z. Xiong, and C. N. Georghiades, "Compression of binary sources with side information at the decoder using LDPC codes," *IEEE Communications Letters*, vol. 6, no. 10, pp. 440–442, October 2002.

[5] J. Garcia-Frias and Y. Zhao, "Compression of correlated binary sources using turbo codes," *IEEE Communications Letters*, vol. 5, no. 10, pp. 417–419, October 2001.

[6] Z. Xiong, A. Liveris, S. Cheng, and Z. Liu, "Nested quantization and Slepian-Wolf coding: a Wyner-Ziv coding paradigm for i.i.d. sources," in *IEEE Workshop on Statistical Signal Processing*, September 2003, pp. 399–402.

[7] S. Pradhan and K. Ramchandran, "Generalized coset codes for distributed binning," *IEEE Trans. on Information Theory*, vol. 51, no. 10, pp. 3457–3474, October 2005.

[8] T. J. Flynn and R. M. Gray, "Encoding of correlated observations," *IEEE Trans. on Information Theory*, vol. 33, no. 6, pp. 773–787, November 1987.

[9] W. Lam and A. R. Reibman, "Design of quantizers for decentralized estimation systems," *IEEE Trans. on Communications*, vol. 41, no. 11, pp. 1602–1605, November 1993.

[10] D. Rebollo-Monedero, R. Zhang, and B. Girod, "Design of optimal quantizers for distributed source coding," in *Proceedings IEEE Data Compression Conference*, March 2003, pp. 13–22.

[11] E. Tuncel, "Predictive coding of correlated sources," in *IEEE Information Theory Workshop*, October 2004, pp. 111–116.

[12] M. Fleming, Q. Zhao, and M. Effros, "Network vector quantization," *IEEE Trans. on Information Theory*, vol. 50, no. 8, pp. 1584–1604, August 2004.

[13] A. Saxena and K. Rose, "Distributed predictive coding for spatio-temporally correlated sources," in *Proceedings IEEE Int. Symp. Information Theory*, June 2007, pp. 1506–1510.

[14] B. Liu and B. Chen, "Channel-optimized quantizers for decentralized detection in sensor networks," *IEEE Trans. on Information Theory*, vol. 52, no. 7, pp. 3349–3358, July 2006.

[15] A. Saxena, J. Nayak, and K. Rose, "On efficient quantizer design for robust distributed source coding," in *Proceedings IEEE Data Compression Conference*, March 2006, pp. 63–71.

[16] J. Karlsson, N. Wernersson, and M. Skoglund, "Distributed scalar quantizers for noisy channels," in *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, April 2007, pp. 633–636.

[17] N. Wernersson, J. Karlsson, and M. Skoglund, "Distributed scalar quantizers for Gaussian channels," in *Proceedings IEEE Int. Symp. Information Theory*, June 2007, pp. 1741–1745.

[18] N. Farvardin and V. Vaishampayan, "Optimal quantizer design for noisy channels: An approach to combined source–channel coding," *IEEE Trans. on Information Theory*, vol. 33, no. 6, pp. 827–838, November 1987.

[19] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*. Dordrecht, The Netherlands: Kluwer academic publishers, 1992.

[20] J. G. Proakis, *Digital Communications*, 4th ed. McGraw-Hill, 2001.

[21] V. Vaishampayan and N. Farvardin, "Joint design of block source codes and modulation signal sets," *IEEE Trans. on Information Theory*, vol. 38, no. 4, pp. 1230–1248, July 1992.

[22] P. Knagenhjelm, "A recursive design method for robust vector quantization," in *Proc. Int. Conf. on Signal Processing Applications and Technology*, November 1992, pp. 948–954.

[23] S. Gadkari and K. Rose, "Noisy channel relaxation for VQ design," in *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 1996, pp. 2048–2051.

[24] A. Fuldseth and T. A. Ramstad, "Bandwidth compression for continuous amplitude channels based on vector approximation to a continuous subset of the source signal space," in *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Munich, Germany, April 1997, pp. 3093–3096.

[25] A. B. Wagner, S. Tavildar, and P. Viswanath, "Rate region of the quadratic Gaussian two-encoder source-coding problem," *IEEE Trans. on Information Theory*, vol. 54, no. 5, pp. 1938–1961, May 2009.

[26] J. J. Xiao and Z. Q. Luo, "Multiterminal source-channel communication over an orthogonal multiple-access channel," *IEEE Trans. on Information Theory*, vol. 53, no. 9, pp. 3255–3264, September 2007.

[27] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Wiley-Interscience, 1991.

[28] A. Mehes and K. Zeger, "Binary lattice vector quantization with linear block codes and affine index assignments," *IEEE Trans. on Information Theory*, vol. 44, no. 1, pp. 79–94, January 1998.

[29] C. E. Shannon, "Communication in the presence of noise," *Proc. IRE*, pp. 10–21, January 1949.

[30] V. Vaishampayan and S. I. R. Costa, "Curves on a sphere, shift-map dynamics, and error control for continuous alphabet sources," *IEEE Trans. on Information Theory*, vol. 49, no. 7, pp. 1658–1672, July 2003.

[31] N. Wernersson, M. Skoglund, and T. Ramstad, "Analog source-channel codes based on orthogonal polynomials," in *Asilomar Conference on Signals, Systems and Computers*, November 2007.

[32] N. Wernersson and M. Skoglund, "Nonlinear coding and estimation for correlated data in wireless sensor networks," *IEEE Trans. on Communications*, vol. 57, no. 10, pp. 2932–2939, October 2009.

Paper B

Low-Delay Joint Source–Channel Mappings for the Gaussian MAC

Johannes Kron, Fady Alajaji, and Mikael Skoglund

Submitted to *IEEE Transactions on Communications*.

Low-Delay Joint Source–Channel Mappings for the Gaussian MAC

Johannes Kron, Fady Alajaji, and Mikael Skoglund

Abstract

In this paper, the bivariate Gaussian multiterminal source coding problem with transmission over the Gaussian multiple-access channel is studied. We propose the use of low-delay joint source–channel mappings and show for the matched bandwidth case how performance saturation, which is unavoidable with linear transmission, can be overcome by optimizing the mappings. To gain deeper understanding, the resulting mappings and their corresponding joint decoders are visualized. The optimized mappings are in general nonlinear and perform a combination of signaling and transmission of analog data. We next look at bandwidth expansion and optimize mappings based on M-ary modulation. The implementation aspects of the design algorithm are also discussed in detail.

1 Introduction

For the well-known CEO problem [1, 2] and the Gaussian multiple-access channel (MAC) with matched bandwidth, it is known that uncoded, or linear, transmission is optimal [3]. Based on this conclusion, it is sometimes wrongly assumed that linear transmission is a good choice in other similar cases, especially if there are low-delay constraints. However, if the CEO problem is modified slightly such that the transmissions are made over orthogonal[1] channels (e.g., orthogonal in time or frequency), linear transmission is no longer optimal because not all of the available channel space is used. In this case, linear transmission is often suboptimal even if one only considers low-delay systems that are uncoded in the sense that

[1]By *orthogonal* we mean that the channels are such that there is no interference between them and that their noise terms are independent.

transmissions are made on a sample-by-sample basis. In general, there exist nonlinear mappings that outperform linear transmission [4].

In this paper we study the bivariate Gaussian multiterminal source coding problem with transmission over the Gaussian MAC. In the CEO problem, the terminals have access to noisy observations of a single underlying source variable that should be estimated. The multiterminal source coding problem differs from the CEO problem in that, here, the source variable at *each* terminal should be estimated. In this scenario with matched bandwidth, it has been shown that linear transmission is optimal only for signal-to-noise ratios (SNRs) below a certain threshold determined by the correlation between the sources [5]. In [5], a joint source–channel code was proposed that is shown to outperform separation-based schemes [5,6]. However, the results are asymptotic in the sense that infinite block lengths are assumed. Motivated by low-delay constraints, in for example sensor networks or closed-loop control applications, we look at low-dimensional joint source–channel mappings.

Low-dimensional joint source–channel coding is a well-studied topic for point-to-point communication systems. Scalar quantization with transmission over noisy channels was studied in [7], where different index assignments were compared. An alternative approach to optimizing the index assignment is to take the channel into consideration when the quantization regions are designed. This is the strategy of [8], where a generalization of the Lloyd–Max algorithm is used to find channel-optimized quantizers. Generalizations to vector sources were done in [9–12]. The physical channels that we use for transmission are analog and not digital. When we talk about digital channels, we implicitly assume that a mapping from our digital channel to the analog physical channel has been provided. The energy efficiency of the communication system could be increased if the source symbols are mapped directly to the analog channel. The idea of this kind of analog codes was mentioned already by Shannon in [13]. Theoretical characterization of optimal analog communication systems can be found in [14–16]. If there is a mismatch between the channel and source dimensions, such that their ratio is greater than or less than one, the system is usually referred to as bandwidth expansion or bandwidth compression, respectively. If the ratio is exactly one, we refer to the system as matched bandwidth. One important observation is that a linear system is not optimal in general. In the case of bandwidth compression, a linear system would need to discard some of the dimensions of the source. On the other hand for bandwidth expansion, a linear system would only use a subspace of the channel space, which means that it does not use all of the available degrees of freedom. Practical results of *point-to-point communication* with this kind of analog source–channel mappings can be found in [17–24] for the case of bandwidth compression and [19, 21–28] for the bandwidth expansion case. There are not as many results for multi-user scenarios as considered in this paper;

see [4, 29–32] for a few of the existing multi-user results.

In what follows, we explicitly design *optimized, low-delay* transmission schemes for the Gaussian multiterminal source coding problem and the Gaussian MAC. This problem was also considered recently in [32], however using parametrized mappings, and in [29, 31], assuming orthogonal channels. The novel contribution of this paper is to consider numerically optimized mappings for the *interfering* Gaussian MAC. The outline of the paper is as follows: In Section 2, we introduce the problem and define the constraints on the system. In Section 3, we propose an iterative design algorithm that can be used to find numerically optimized solutions. We present lower bounds and the performance of linear transmission in Section 4. Section 5 deals with implementation aspects and in Section 6, we present our numerical results from using the design algorithm. Finally, we conclude the paper in Section 7.

2 Problem Formulation

We consider a scenario where two spatially separated sensor nodes, each measure a Gaussian random variable X_i, which is to be transmitted to a joint receiver. The Gaussian random variables, X_1 and X_2, are identically distributed and have zero mean and variance σ_X^2. Furthermore, they are correlated with a correlation coefficient defined by

$$\rho \triangleq \frac{E[X_1 X_2]}{\sigma_X^2}. \tag{1}$$

Due to low-delay constraints, we consider transmission on a sample-by-sample basis where sensor i maps its measurement directly to the channel space by the mapping $\boldsymbol{\alpha}_i$ and transmits $\boldsymbol{S}_i = \boldsymbol{\alpha}_i(X_i) \in \mathcal{S}^N \subseteq \mathbb{R}^N$. The set $\mathcal{S}$ consists of all possible modulation points that can be used by the sensor nodes in each channel use; it is given by

$$\mathcal{S} = \left\{ -\Delta\frac{M-1}{2}, -\Delta\frac{M-3}{2}, \ldots, \Delta\frac{M-3}{2}, \Delta\frac{M-1}{2} \right\}, \tag{2}$$

where Δ determines the resolution of the modulation points and M is the number of points, or the cardinality, of $\mathcal{S}$. For finite values of M, this is the same as pulse-amplitude modulation (PAM) and as M and Δ approach infinity and zero, respectively, we have analog modulation[2]. N is a bandwidth expansion factor, if $N = 1$ we have a matched bandwidth between the source and channel, whereas if $N > 1$ we have bandwidth expansion.

[2]Of course one has to make sure that $1/\Delta$ does not grow faster than M in which case the product $\Delta M \to 0$.

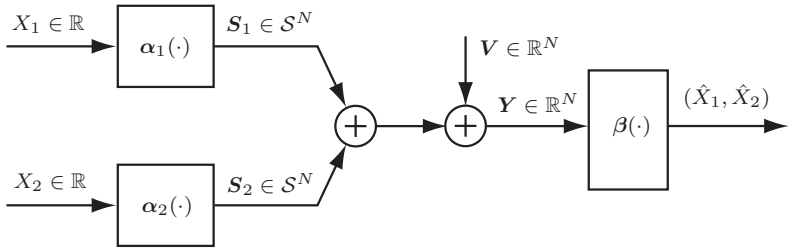


Figure 1: Overview of the joint source–channel coding scheme for the Gaussian MAC with correlated sources.

The transmissions are made over the Gaussian MAC with average power constraint P and the signal received at the receiver is given by

$$\boldsymbol{Y} = \boldsymbol{\alpha}_1(X_1) + \boldsymbol{\alpha}_2(X_2) + \boldsymbol{V}, \tag{3}$$

as shown in Fig. 1, where $\boldsymbol{V}$ is additive white Gaussian noise, independent of X_1 and X_2, with covariance matrix $\sigma_V^2 \boldsymbol{I}$. At the receiver, the source variables (X_1, X_2) are reconstructed by the mapping $\boldsymbol{\beta}(\boldsymbol{Y}) = (\beta_1(\boldsymbol{Y}), \beta_2(\boldsymbol{Y}))$, where

$$\hat{X}_i = \beta_i(\boldsymbol{Y}), \qquad i = 1, 2. \tag{4}$$

The overall objective is to find the combination of $(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \beta_1, \beta_2)$ that minimizes the average mean squared error, defined according to

$$\text{MSE} \triangleq \frac{1}{2} \sum_{i=1}^{2} E[(X_i - \hat{X}_i)^2]. \tag{5}$$

The minimization should be done under the following transmission power constraint:

$$\frac{1}{2} \sum_{i=1}^{2} E[\|\boldsymbol{\alpha}_i(X_i)\|^2] \leq P. \tag{6}$$

3 Analysis

Using the Lagrange multiplier method [33], we turn the constrained optimization problem of minimizing (5) subject to (6) into an unconstrained problem by first forming the Lagrange cost function

$$J(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \beta_1, \beta_2) = \sum_{i=1}^{2} \left\{ \frac{1}{2} E[(X_i - \hat{X}_i)^2] + \lambda E[\|\boldsymbol{\alpha}_i(X_i)\|^2] \right\} \tag{7}$$

and next expressing the problem as

$$\min_{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\beta_1,\beta_2} J(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\beta_1,\beta_2). \tag{8}$$

Here λ is a Lagrange multiplier that is used to control the average power. If for a given λ, we solve the unconstrained problem in (8) and find that the power constraint in (6) is fulfilled with equality, the solution we have obtained is also a solution to the constrained optimization problem [33]. Another way of putting it is that we pick a λ and solve (8). For a particular choice of P, the inequality in (6) will be tight, consequently, this is the corresponding constrained optimization problem that was solved with our choice of λ.

The minimization problem in (8) is still very hard to solve due to interdependencies between the components that are optimized and since the problem is nonconvex. To get around these difficulties we proceed as in vector quantization (VQ) optimization [34, 35] and state necessary conditions for the optimality of each component. Based on these necessary conditions, we optimize the system iteratively, one component at a time while keeping the other components fixed.

3.1 Necessary Conditions for Optimality

Beginning with $\boldsymbol{\alpha}_1$, if we assume that $(\boldsymbol{\alpha}_2,\beta_1,\beta_2)$ are fixed, the optimal $\boldsymbol{\alpha}_1$ is given by

$$\begin{aligned}\boldsymbol{\alpha}_1 &= \arg\min_{\boldsymbol{\alpha}_1} J(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\beta_1,\beta_2)\\ &= \arg\min_{\boldsymbol{\alpha}_1}\Big\{\sum_{i=1}^{2}\Big(\frac{1}{2}\underbrace{E[(X_i-\hat{X}_i)^2]}_{\mathrm{MSE}_i}\Big)+\lambda\underbrace{E[\|\boldsymbol{\alpha}_1(X_1)\|^2]}_{P_1}\Big\}.\end{aligned} \tag{9}$$

Using Bayes' rule, the MSE of user i can be expressed as

$$\begin{aligned}\mathrm{MSE}_i = &\int p(x_1)\int p(x_2|x_1)\times\\ &\int p(\boldsymbol{y}|\boldsymbol{\alpha}_1(x_1),\boldsymbol{\alpha}_2(x_2))(x_i-\beta_i(\boldsymbol{y}))^2\mathrm{d}\boldsymbol{y}\mathrm{d}x_2\mathrm{d}x_1,\end{aligned} \tag{10}$$

where $p(\cdot)$ and $p(\cdot|\cdot)$ denote probability densities and conditional densities, respectively. The average power of user i is given by

$$P_i = \int p(x_i)\|\boldsymbol{\alpha}_i(x_i)\|^2\mathrm{d}x_i. \tag{11}$$

Looking at (10) and (11), we can see that the objective function in (9) can be expressed as an integration over x_1 with an integrand of the form

$p(x_1)f(x_1)$. Since $p(x_1)$ by definition is nonnegative for all x_1, we can find the optimal $\boldsymbol{\alpha}_1$ by minimizing $f(x_1)$ for each value of x_1 in the following way:

$$\boldsymbol{\alpha}_1(x_1) = \arg\min_{s_1 \in \mathcal{S}^N} \Big\{ \int p(x_2|x_1) \int p(\boldsymbol{y}|\boldsymbol{s_1}, \boldsymbol{\alpha}_2(x_2)) \times \frac{1}{2}\Big((x_1 - \beta_1(\boldsymbol{y}))^2 + (x_2 - \beta_2(\boldsymbol{y}))^2\Big) \mathrm{d}\boldsymbol{y}\mathrm{d}x_2 + \lambda \|s_1\|^2 \Big\}. \quad (12)$$

The condition as given in (12) is a necessary condition for $\boldsymbol{\alpha}_1$ to be part of the optimal solution and if $(\boldsymbol{\alpha}_2, \beta_1, \beta_2)$ are given beforehand it provides the optimal mapping $\boldsymbol{\alpha}_1$. Obviously, $\boldsymbol{\alpha}_2$ can be found in a similar way by assuming that $(\boldsymbol{\alpha}_1, \beta_1, \beta_2)$ are fixed.

If we move on to the receiving side and assume that $\boldsymbol{\alpha}_1$ and $\boldsymbol{\alpha}_2$ are given, we can find the optimal estimators. Since we are using the MSE as our cost function, β_i is given by the conditional expected mean

$$\hat{x}_i = \beta_i(\boldsymbol{y}) = E[X_i|\boldsymbol{y}]. \quad (13)$$

3.2 Design Algorithm

Based on the necessary conditions for optimality, we now propose a design algorithm that iterates between optimizing the mappings at the sensor nodes and the receiver. This kind of iterative optimization does not in general guarantee convergence to the global optimum. Nevertheless, it has successfully been used in many applications such as VQ design [34, 35] and design of joint source–channel mappings for different scenarios, see for example [8, 18, 30]. For a fixed SNR $\triangleq P/(N\sigma_V^2)$ and correlation ρ, the design procedure is stated in Algorithm 1.

In our simulations we have used $\delta = 10^{-4}$ to determine when to stop the iteration. It is clear that the algorithm will converge since the cost function is bounded from below and reduced by each iteration. The question is how to avoid poor local minimums. In this paper we use a combination of two methods. We first choose a good initialization in the first step of the design algorithm, see Section 6.1. We also use a post-processing step similar to noisy channel relaxation [36]. Assume that we have designed systems for a range of SNR points, for example $\{10, 15, 20, 25\}$ dB. We now initialize the algorithm with the system designed for the smallest SNR and optimize for the second smallest SNR. If the new system performs better than the system that was previously optimized for this SNR point we keep it. This process is repeated for each SNR point. Once we reach the largest SNR point we repeat the process backwards until we reach the smallest SNR. By stepping back and forward a couple of times, we can avoid that a system designed for a particular SNR performs much worse than any other system. This process was also used in [8].

Algorithm 1 Design Algorithm

Require: Initial mappings of $\boldsymbol{\alpha}_1$ and $\boldsymbol{\alpha}_2$, the SNR and correlation for which the system should be optimized and the threshold δ that determines when to stop the iterations.

Ensure: Locally optimized $(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \beta_1, \beta_2)$.

1: Find the optimal receivers (β_1, β_2) by using (13).
2: Set the iteration index $k = 0$ and $J^{(0)} = \infty$.
3: **repeat**
4: Set $k = k + 1$
5: Find the optimal source mapping $\boldsymbol{\alpha}_1$ by using (12).
6: Find the optimal receivers (β_1, β_2) by using (13).
7: Find the optimal source mapping $\boldsymbol{\alpha}_2$ by using (12).
8: Find the optimal receivers (β_1, β_2) by using (13).
9: Evaluate the cost function $J^{(i)}$ according to (7).
10: **until** $(J^{(k-1)} - J^{(k)})/J^{(k-1)} < \delta$

4 Lower Bounds and Linear Transmission

The distortion region of this problem is in general unknown. Since the sources are correlated and the communication is over the GMAC, the source–channel separation theorem does not hold. We can lower bound the distortion by considering collaboration between the sensor nodes such that blocks of source samples X_1 and X_2 are encoded jointly. By letting the block length tend to infinity, we have a point-to-point communications problem where the source–channel separation theorem *does* hold. Consequently, we can find a lower bound on the distortion by evaluating the distortion-rate function of an i.i.d. Gaussian source $\boldsymbol{X} = (X_1, X_2)^T$ at a rate R equal to N times the capacity of the channel at hand. The general distortion-rate function in [37, Ch. 4] can be specialized to our two-dimensional source $\boldsymbol{X}$ and is then given by

$$D(R) = \begin{cases} \sigma_X^2 \dfrac{1 - |\rho| + (1 + |\rho|)2^{-2R}}{2}, & \text{if } R \leq \dfrac{1}{2}\log_2 \dfrac{1 + |\rho|}{1 - |\rho|}, \\ \sigma_X^2 \sqrt{1 - \rho^2}\, 2^{-R}, & \text{if } R > \dfrac{1}{2}\log_2 \dfrac{1 + |\rho|}{1 - |\rho|}. \end{cases} \tag{14}$$

Here $D(R)$ is a lower bound on the average MSE, as defined in (5), if the source is encoded with R bits per source vector $\boldsymbol{X}$. For the matched bandwidth case where we have $\mathcal{S} \approx \mathbb{R}$, we set R to the capacity of the AWGN channel. In the cases where we have bandwidth expansion and $\mathcal{S}$ consists of a small number of constellation points, we set R to N times the capacity

of the constellation-constrained AWGN channel[3]. It is also important to realize that since the output of the virtual joint encoder comes from two sensor nodes that combine their outputs coherently, we have a 3 dB power gain in our point-to-point channel in comparison to an ordinary channel. To compensate for this, the capacity of the point-to-point channel is evaluated at an equivalent SNR of $2P/(N\sigma_V^2)$.

Analog Modulation and Matched Bandwidth

In the special case of $\mathcal{S} = \mathbb{R}$ and $N = 1$, we can compare the performance of our optimized systems to linear transmission and the lower bound from [5]. For the reader's convenience we include the result for the symmetric case, that is, $\text{MSE}_1 = \text{MSE}_2$. Linear transmission is one of the most simple transmission schemes, where the outputs of $\boldsymbol{\alpha}_1$ and $\boldsymbol{\alpha}_2$ are scaled versions of their inputs. Given the power constraint in (6), the optimal scaling is such that the average power of each transmitter is P. The MSE of linear transmission is given by

$$\text{MSE}_{\text{linear}} = \sigma_X^2 \frac{P(1-\rho^2)+\sigma_V^2}{2P(1+\rho)+\sigma_V^2} \tag{15}$$

and a lower bound on the MSE is given by

$$\text{MSE}_{\text{lower}} = \begin{cases} \sigma_X^2 \frac{P(1-\rho^2)+\sigma_V^2}{2P(1+\rho)+\sigma_V^2} & \text{for } P/\sigma_V^2 \le \rho/(1-\rho^2) \\ \sigma_X^2 \sqrt{\frac{(1-\rho^2)\sigma_V^2}{2P(1+\rho)+\sigma_V^2}} & \text{for } P/\sigma_V^2 > \rho/(1-\rho^2). \end{cases} \tag{16}$$

As can be seen by comparing (15) and (16), the lower bound is tight for $\text{SNR} = P/\sigma_V^2 \le \rho/(1-\rho^2)$ in which case linear transmission is optimal. The intuition why linear transmission is optimal only below this threshold, is that for low SNRs the power of the two sensor nodes combine coherently to combat the channel noise, but as the SNR increases, the difference between the two signals will instead cause interference that is greater than the channel noise itself. In [5], a joint source–channel code is constructed that performs close to the lower bound for high SNRs. However, both the lower bound and the joint source–channel code rely on infinite block lengths. Note that schemes based on source–channel separation are in general suboptimal if the source samples are correlated.

5 Implementation Aspects

With no loss of generality we assume that the transmit power P is always one and that the $\text{SNR} = P/(N\sigma_V^2)$ is varied by changing the noise variance σ_V^2.

[3]In general, there are no closed-form expressions in this case and one has to find the capacity by means of numerical integrations.

For the actual implementation of the formulas in (12) and (13), we need to make some modifications and approximations. First of all, we cannot evaluate the formulas for all (x_1, x_2) in the real domain. Instead we generate Monte-Carlo samples from the distribution of X_1 and assign them to the set $\mathcal{X}$. The basic idea is to let the samples in this set represent both X_1 and X_2 by associating each element $x_1 \in \mathcal{X}$ with a new set $\mathcal{X}_{2|x_1} \subseteq \mathcal{X}$, where each element is "typically" correlated with x_1. The elements in $\mathcal{X}_{2|x_1}$ are randomly chosen among the samples in the set

$$\mathcal{X}_{\gamma,x_1} = \{x_2 : x_2 \in \mathcal{X}, p(x_2|x_1) > \gamma\}, \tag{17}$$

where

$$p(x_2|x_1) = \frac{1}{\sqrt{2\pi(1-\rho^2)\sigma_X^2}} \exp\Big(- \frac{(x_2 - \rho x_1)^2}{2(1-\rho^2)\sigma_X^2}\Big), \tag{18}$$

which are those samples in $\mathcal{X}$ whose conditional probability density function given x_1 are higher than the threshold γ. In our simulations we have used 5000 samples to define $\mathcal{X}$ and $\gamma = 0.001$ as a tradeoff between complexity and accuracy. The idea of choosing samples randomly and not including all samples from this set is to further reduce the complexity. The acceptance ratio has been set so that the cardinality of $\mathcal{X}_{2|x_1}$ is around 400. We need to shape $p(x_2|x_1)$ slightly to compensate for the fact that the samples are drawn from an already Gaussian distribution instead of a uniform distribution and also since we have cut off the tails of the distribution with our threshold γ. This is done by letting

$$P(x_2|x_1) = \frac{p(x_2|x_1)}{\kappa_{x_1} p(x_2)}, \tag{19}$$

where κ_{x_1} is a normalization constant such that

$$\sum_{x_2 \in \mathcal{X}_{2|x_1}} P(x_2|x_1) = 1, \qquad \forall x_1 \in \mathcal{X}. \tag{20}$$

Our second problem is that the received signal still is real-valued. To get around this, the input to the receiver is pre-quantized by a nearest neighbor quantizer to a finite set $\mathcal{Y}$ defined by

$$\mathcal{Y} = \left\{ -\frac{\Delta}{\eta}\frac{\tilde{M}-1}{2}, -\frac{\Delta}{\eta}\frac{\tilde{M}-3}{2}, \ldots, \frac{\Delta}{\eta}\frac{\tilde{M}-3}{2}, \frac{\Delta}{\eta}\frac{\tilde{M}-1}{2} \right\}, \tag{21}$$

where $\eta \in \mathbb{N}$ is a nesting factor and $\tilde{M}$ is chosen such that $y = s_1 + s_2 \in \mathcal{Y}$ for all $s_1, s_2 \in \mathcal{S}$. By using a nesting factor $\eta > 1$, the quantization at the receiver has a finer resolution and thus provides soft information of the transmitted symbols instead of making hard decisions.

We are now in position to state the discretized versions of (12) and (13), which are used in the actual implementation of the design algorithm. They are given by

$$\boldsymbol{\alpha}_1(x_1) = \arg\min_{\boldsymbol{s}_1 \in \mathcal{S}^N} \Big\{ \sum_{i=1}^{2} \sum_{x_2 \in \mathcal{X}_{2|x_1}} P(x_2|x_1) \times \sum_{\boldsymbol{y} \in \mathcal{Y}^N} P(\boldsymbol{y}|\boldsymbol{s}_1, \boldsymbol{\alpha}_2(x_2))(x_i - \beta_i(\boldsymbol{y}))^2 + \lambda \|\boldsymbol{s}_1\| 2 \Big\} \tag{22}$$

and

$$\beta_1(\boldsymbol{y}) = \frac{\sum_{x_1 \in \mathcal{X}} x_1 \sum_{x_2 \in \mathcal{X}_{2|x_1}} P(x_2|x_1) P(\boldsymbol{y}|\boldsymbol{\alpha}_1(x_1), \boldsymbol{\alpha}_2(x_2))}{\sum_{x_1 \in \mathcal{X}} \sum_{x_2 \in \mathcal{X}_{2|x_1}} P(x_2|x_1) P(\boldsymbol{y}|\boldsymbol{\alpha}_1(x_1), \boldsymbol{\alpha}_2(x_2))}. \tag{23}$$

The expressions for $\boldsymbol{\alpha}_2$ and β_2 follow by straightforward modifications of the above formulas.

6 Performance Evaluation

We consider three different sets of modulation points, which we denote by analog, 16-QAM, and BPSK modulation. In the first case we let $N = 1$ and approximate analog modulation by letting Δ be small in relation to the standard deviation of the channel noise and M large (i.e., $\mathcal{S} \approx \mathbb{R}$). To be more specific, we have kept $\Delta(M-1)/2 = 4$ and used an M in the range $[321, 1281]$ depending on the SNR. For the 16-QAM case we let $N = 2$ and use $M = 4$ modulation points per channel dimension and for the BPSK case we let $N = 4$ and use $M = 2$ modulation points per channel dimension. In the latter two cases, we have used a Δ such that (6) is fulfilled with equality if all modulation points are equiprobable. The nesting factor η that is used on the receiving side to have a finer resolution of the constellations points is set to 1 in the analog scenario and 2 in the 16-QAM and BPSK scenarios.

6.1 Initialization

Each scenario requires its own initialization; below we describe the initializations that have been used in the design algorithm for each scenario.

Analog ($N = 1$)

After running the design algorithm a couple of times, it became clear that staircase functions are good starting points for the iterative design procedure

in the analog case. We therefore initialize α_1 with a staircase function with L steps. Mathematically, this can be expressed as

$$\alpha_1(x_1) = \left\lfloor Q(x_1) \frac{2 \max\limits_{s_1 \in \mathcal{S}} s_1}{L-1} \right\rceil_{\mathcal{S}}, \tag{24}$$

where

$$Q(x_1) = \left\lfloor x_1 \frac{L-1}{2 \max\limits_{x_1 \in \mathcal{X}} |x_1|} - \frac{\text{even}(L)}{2} \right\rceil_{\mathbb{Z}} + \frac{\text{even}(L)}{2}, \tag{25}$$

$$\text{even}(L) = \begin{cases} 1 & \text{if } L \text{ is even} \\ 0 & \text{otherwise,} \end{cases} \tag{26}$$

and $\lfloor \cdot \rceil_{\mathcal{S}}$ returns the closest point (in terms of Euclidean distance) in the set $\mathcal{S}$. The best choice of L typically depends on both the correlation and the SNR. In our simulations we have used L in the range from 2 to 20. The mapping α_2 has been initialized identical to zero for all inputs.

16-QAM ($N = 2, M = 4$)

For 16-QAM, we consider two different initializations based on Lloyd–Max quantization [34, 35]. The first strategy is based on quantization with 4 levels at each source node and a mapping of the quantization indices to modulation points using time-division multiple access (TDMA). That is, each source node only uses one dimension of the channel and there is therefore no interference. The second strategy uses identical mappings at the source nodes (i.e., $\boldsymbol{\alpha}_1 = \boldsymbol{\alpha}_2$) and is based on quantization with 16 levels, where the indices are mapped to the channel space in a spiral-like way such that neighboring quantization levels correspond to neighboring modulation points [38].

BPSK ($N = 4, M = 2$)

For BPSK, we initialize in two different ways similarly as in the 16-QAM case. For the TDMA initialization we first quantize with 4 levels and map the quantization indices to modulation points such that there is no interference between the two source encoders. For the second strategy we once again use identical mappings at the source nodes based on quantization with 16 levels. For the mapping of quantization indices to binary modulation points we use the folded binary code [7], which is known to perform well for Gaussian sources over noisy channels.

6.2 Numerical Results

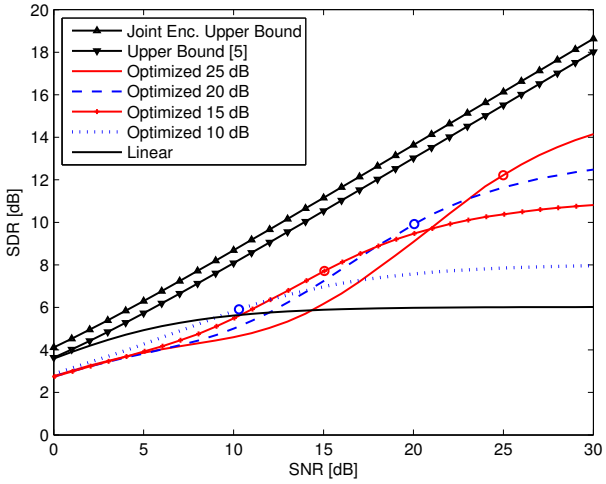
In Figs. 2–4, we plot the results for each set of modulation points and two different source correlations, namely, $\rho = 0.5$ and $\rho = 0.9$. The results are plotted as SDR $\triangleq 10 \log_{10} \sigma_X^2 / \text{MSE}$ versus SNR. The curves show practical systems where the encoders have been optimized for certain SNR points that are marked by circles in the figures. We assume that the true SNR is known by the receiver and that the decoder is updated accordingly.

Analog ($N = 1$)

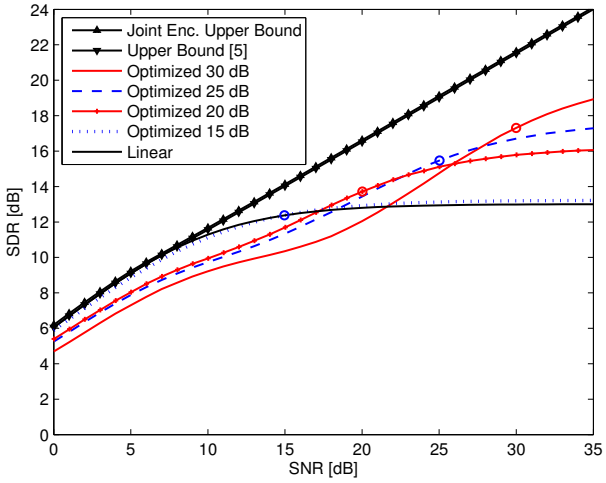
Starting with the results of analog modulation in Fig. 2, we can see that the optimized systems perform very well and overcome the saturation that is unavoidable with linear transmission. The systems are robust against SNR mismatch on the transmitting side. Looking, for example, at the system optimized for SNR = 15 dB and $\rho = 0.5$ in Fig. 2(a), we can see that it performs very well in the range from 10 dB to 20 dB. In a real system where there are time variations, a feedback link could be utilized to ensure that the best performing encoders are always used. Since we are plotting the results as SDR, the distortion lower bounds in Section 4 become SDR upper bounds. It should be emphasized that the upper bounds are asymptotic results in the sense of infinite block lengths of the source samples as well as in the channel coding part. The gap between the upper bounds and our low-delay system is therefore not surprising. As can be seen, there is a gap between the two upper bounds in the medium correlation case ($\rho = 0.5$) but as the correlation increases, the two bounds coincide. This is because the upper bound based on joint encoding always has a 3 dB power gain due to coherent combining of the signals regardless of the correlation. Intuitively, this is too optimistic since, for example, if the source correlation would be zero there would be no power gain at all due to coherent combining. The upper bound in [5] takes this dependency into account and is therefore a tighter bound. Nevertheless, the upper bound based on joint encoding is useful in the bandwidth expansion scenarios (i.e., $N > 1$) that are evaluated next.

16-QAM and BPSK modulation

Looking at the performance of the 16-QAM and BPSK systems in Figs. 3–4, the gap to the upper bound seems to increase as the bandwidth expansion factor (i.e., N) increases. The same phenomenon has been observed in other cases where low-delay joint source–channel mappings for bandwidth expansion have been studied [22, 30]. As methods of reference, we have used the initialization systems based on TDMA and identical mappings, as described in Section 6.1, with the corresponding *optimal* receiver based on full channel-state information. Using identical mappings is comparable

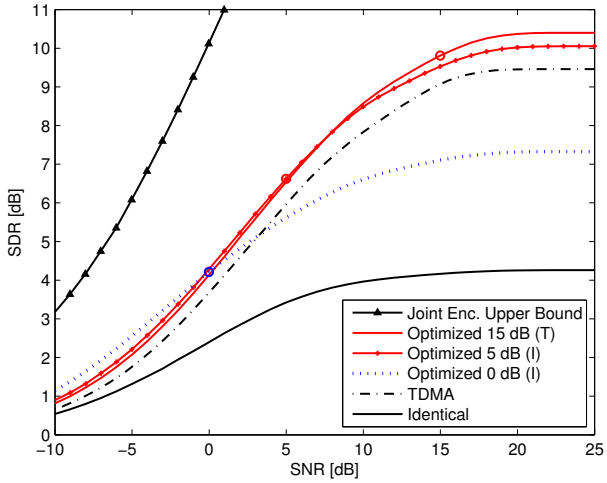


(a) $N = 1, \rho = 0.5$

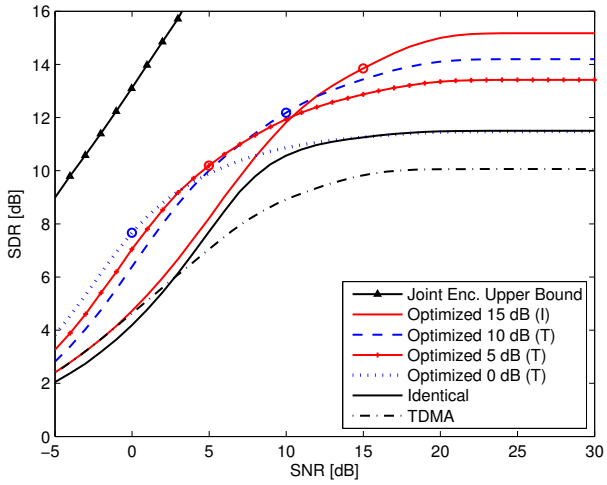


(b) $N = 1, \rho = 0.9$

Figure 2: Simulation results for analog modulation with matched bandwidth. Circles mark the points where the mappings are optimized.

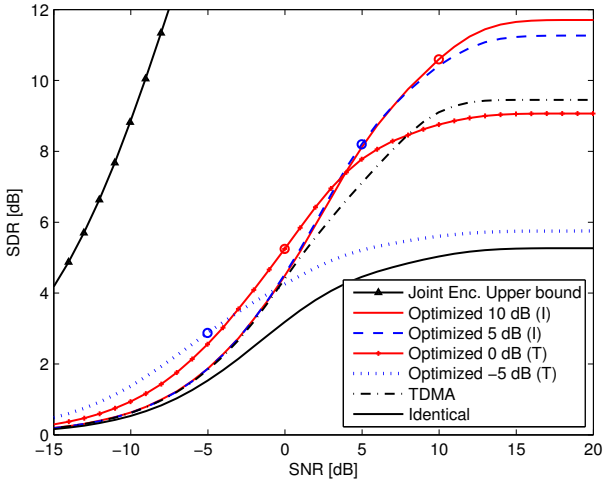


(a) $N = 2, \rho = 0.5$

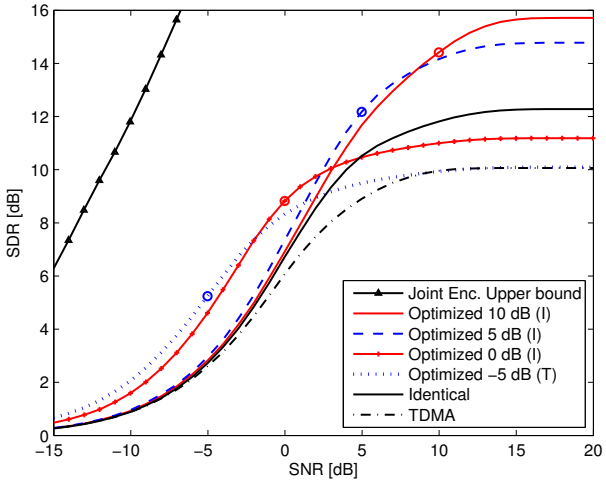


(b) $N = 2, \rho = 0.9$

Figure 3: Simulation results for 16-QAM with bandwidth expansion. Circles mark the points where the mappings are optimized. The letters T and I, corresponding to TDMA and identical, are used to indicate which initialization that was used for each optimized system.



(a) $N = 4, \rho = 0.5$



(b) $N = 4, \rho = 0.9$

Figure 4: Simulation results for BPSK modulation with bandwidth expansion. Circles mark the points where the mappings are optimized. The letters T and I, corresponding to TDMA and identical, are used to indicate which initialization that was used for each optimized system.

to linear transmission in the analog case, which is one of the reasons this is a suitable reference for the bandwidth expansion cases. In comparison to these systems that have the same complexity and delay, the optimized mappings perform very well and give large performance gains. It is worth to mention that also the upper bound saturates eventually due to the finite number of modulation points

The two different methods of initialization (i.e., TDMA and identical) give similar results after the design algorithm and the post-processing step have been executed. In Figs. 3–4, we show the results of the methods that have the best performance at the SNR points where the systems are optimized. Which initialization that was used for each particular system is indicated by T for TDMA and I for identical in the legends. It is hard to make any general recommendations for which initialization to use; it all depends on the modulation scheme, SNR, and the source correlation.

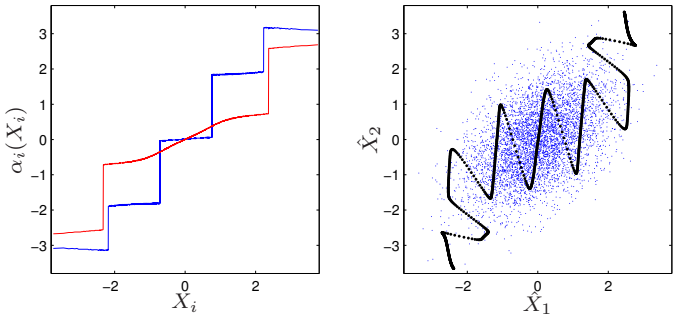
6.3 Encoders and Decoders

We shall now take a closer look at the encoder–decoder structure in the analog scenario. In Fig. 5, we show encoders (left) and decoders (right) for different correlations and SNRs. The interaction between the two encoders is easiest understood by looking at the corresponding joint decoders to the right in Fig. 5. The staircase-like behavior of the encoder mappings performs a combination of signaling and transmission of analog data, which makes the reconstruction points fill the source space in an efficient way. This is the reason why the optimized mappings perform better than linear transmission where the decoder would be the straight line $\hat{X}_2 \equiv \hat{X}_1$ regardless of the SNR and correlation.

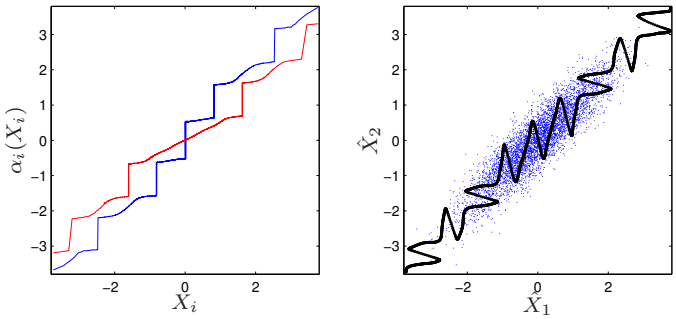
It is a bit harder to visualize the encoders for the bandwidth expansion cases because the outputs are two- and four-dimensional. To give some insights into how the mappings perform, we show decoders for two cases with BPSK modulation in Fig. 6; the mappings are optimized for an SNR of 10 dB and correlations $\rho = 0.5$ and $\rho = 0.9$. We can see that the reconstruction points does not follow a line anymore as in the analog case, the same is also true for the 16-QAM mappings. This is simply because the input to the decoder is a multi-dimensional vector. Both of the mappings in Fig. 6 were initialized using identical mappings.

7 Conclusions

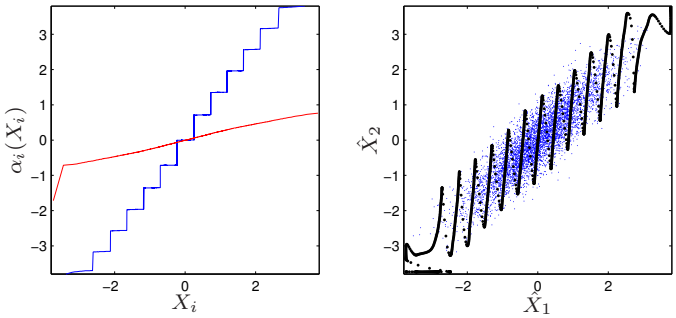
We have proposed the use of optimized source–channel mappings for the bivariate Gaussian multiterminal source coding problem with transmission over the Gaussian MAC. For analog modulation and matched bandwidth, the optimized nonlinear mappings fill the gap between linear transmission



(a) SNR = 15 dB, $\rho = 0.5$



(b) SNR = 20 dB, $\rho = 0.9$



(c) SNR = 30 dB, $\rho = 0.9$

Figure 5: Analog modulation with matched bandwidth: Encoders (left) and their corresponding decoders (right) optimized for different SNRs and correlations. In the figures to the right, the dotted lines show reconstruction points $(\hat{x}_1, \hat{x}_2)$ and the small dots are samples from the distribution of (X_1, X_2).

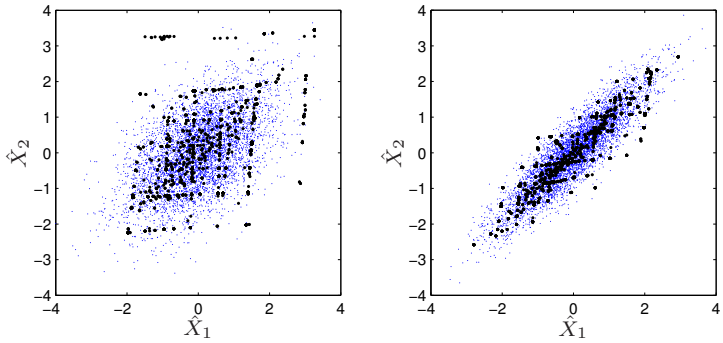


Figure 6: BPSK modulation with bandwidth expansion ($N = 4$): Decoders optimized for an SNR of 10 dB and correlation $\rho = 0.5$ (left) and $\rho = 0.9$ (right). The larger dots show reconstruction points $(\hat{x}_1, \hat{x}_2)$ and the small dots are samples from the distribution of (X_1, X_2).

and existing upper bounds. Also in the case of 16-QAM and BPSK modulation with bandwidth expansion, we have shown that it is possible to have large performance gains by optimizing the mappings. The main advantage of the optimized mappings is their low-delay properties, due to their operation on a sample-by-samples basis.

References

[1] T. Berger, Z. Zhang, and H. Viswanathan, "The CEO problem [multiterminal source coding]," *IEEE Trans. on Information Theory*, vol. 42, no. 3, pp. 887–902, May 1996.

[2] H. Viswanathan and T. Berger, "The quadratic Gaussian CEO problem," *IEEE Trans. on Information Theory*, vol. 43, no. 5, pp. 1549–1559, 1997.

[3] M. Gastpar, "Uncoded transmission is exactly optimal for a simple Gaussian "sensor" network," *IEEE Trans. on Information Theory*, vol. 54, no. 11, pp. 5247–5251, November 2008.

[4] N. Wernersson and M. Skoglund, "Nonlinear coding and estimation for correlated data in wireless sensor networks," *IEEE Trans. on Communications*, vol. 57, no. 10, pp. 2932–2939, October 2009.

[5] A. Lapidoth and S. Tinguely, "Sending a bivariate Gaussian over a Gaussian MAC," *IEEE Trans. on Information Theory*, vol. 56, no. 6, pp. 2714–2752, 2010.

[6] R. Rajesh and V. Sharma, "Source channel coding for Gaussian sources over a Gaussian multiple access channel," in *Proc. 45 Allerton conference on Computing Control and Communication*, 2007.

[7] N. Rydbeck and C. W. Sundberg, "Analysis of digital errors in nonlinear PCM systems," *IEEE Trans. on Communications*, vol. 24, no. 1, pp. 59–65, January 1976.

[8] N. Farvardin and V. Vaishampayan, "Optimal quantizer design for noisy channels: An approach to combined source–channel coding," *IEEE Trans. on Information Theory*, vol. 33, no. 6, pp. 827–838, November 1987.

[9] H. Kumazawa, M. Kasahara, and T. Namekawa, "A construction of vector quantizers for noisy channels," *Electronics and Communications in Japan*, vol. 67, no. 4, pp. 39–47, January 1984.

[10] K. A. Zeger and A. Gersho, "Vector quantizer design for memoryless noisy channels," in *IEEE International Conference on Communications*, Philadelphia, USA, June 1988, pp. 1593–1597.

[11] N. Farvardin, "A study of vector quantization for noisy channels," *IEEE Trans. on Information Theory*, vol. 36, no. 4, pp. 799–809, July 1990.

[12] N. Farvardin and V. Vaishampayan, "On the performance and complexity of channel-optimized vectorquantizers," *IEEE Trans. on Information Theory*, vol. 37, no. 1, pp. 155–160, January 1991.

[13] C. E. Shannon, "Communication in the presence of noise," *Proc. IRE*, pp. 10–21, January 1949.

[14] J. M. Wozencraft and I. M. Jacobs, *Principles of Communication Engineering*. Wiley, 1965.

[15] D. J. Sakrison, *Transmission of Waveforms and Digital Information*. Wiley, 1968.

[16] J. Ziv, "The behavior of analog commnunication systems," *IEEE Trans. on Information Theory*, vol. 16, no. 5, pp. 587–594, September 1970.

[17] V. Vaishampayan, "Combined source–channel coding for bandlimited waveform channels," Ph.D. dissertation, University of Maryland, 1989.

[18] A. Fuldseth and T. A. Ramstad, "Bandwidth compression for continuous amplitude channels based on vector approximation to a continuous subset of the source signal space," in *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Munich, Germany, April 1997, pp. 3093–3096.

[19] A. Fuldseth, "Robust subband video compression for noisy channels with multilevel signaling," Ph.D. dissertation, NTNU Trondheim, 1997.

[20] S. Chung, "On the construction of some capacity-approaching coding schemes," Ph.D. dissertation, MIT, 2000.

[21] T. A. Ramstad, "Shannon mappings for robust communication," *Telektronikk*, vol. 98, no. 1, pp. 114–128, 2002.

[22] F. Hekland, P. A. Floor, and T. A. Ramstad, "Shannon–Kotel'nikov mappings in joint source–channel coding," *IEEE Trans. on Communications*, vol. 57, no. 1, pp. 94–105, January 2009.

[23] E. Akyol, K. Rose, and T. Ramstad, "Optimal mappings for joint source channel coding," in *IEEE Information Theory Workshop (ITW)*, January 2010.

[24] Y. Hu, J. Garcia-Frias, and M. Lamarca, "Analog joint source channel coding using non-linear mappings and MMSE decoding," *IEEE Trans. on Communications*, to appear.

[25] B. Chen and G. W. Wornell, "Analog error-correcting codes based on chaotic dynamical systems," *IEEE Trans. on Communications*, vol. 46, no. 7, pp. 881–890, July 1998.

[26] V. Vaishampayan and S. I. R. Costa, "Curves on a sphere, shift-map dynamics, and error control for continuous alphabet sources," *IEEE Trans. on Information Theory*, vol. 49, no. 7, pp. 1658–1672, July 2003.

[27] X. Cai and J. W. Modestino, "Bandwidth expansion Shannon mapping for analog error-control coding," in *Information Sciences and Systems, 2006 40th Annual Conference on*, March 2006, pp. 1709–1712.

[28] N. Wernersson, M. Skoglund, and T. Ramstad, "Polynomial based analog source–channel codes," *IEEE Trans. on Communications*, vol. 57, no. 9, pp. 2600–2606, September 2009.

[29] N. Wernersson, J. Karlsson, and M. Skoglund, "Distributed quantization over noisy channels," *IEEE Trans. on Communications*, vol. 57, no. 6, pp. 1693–1700, June 2009.

[30] J. Karlsson and M. Skoglund, "Optimized low-delay source–channel–relay mappings," *IEEE Trans. on Communications*, no. 5, pp. 1397–1404, May 2010.

[31] E. Akyol, K. Rose, and T. Ramstad, "Optimized analog mappings for distributed source-channel coding," in *Proceedings IEEE Data Compression Conference*, 2010, pp. 159–168.

[32] P. A. Floor, A. N. Kim, N. Wernersson, T. A. Ramstad, M. Skoglund, and I. Balasingham, "On transmission of two correlated Gaussian memoryless sources over a Gaussian MAC using delay-free mappings," *arXiv:1101.5716v1 [cs.IT]*, 2011, submitted to IEEE Transactions on Communications.

[33] H. Everett III, "Generalized Lagrange multiplier method for solving problems of optimum allocation of resources," *Operations Research*, vol. 11, no. 3, pp. 399–417, 1963.

[34] S. P. Lloyd, "Least Squares Quantization in PCM," *IEEE Trans. on Information Theory*, vol. 28, no. 2, pp. 129–137, March 1982.

[35] Y. Linde, A. Buzo, and R. M. Gray, "An algorithm for vector quantizer design," *IEEE Trans. on Communications*, vol. 28, no. 1, pp. 84–95, January 1980.

[36] S. Gadkari and K. Rose, "Noisy channel relaxation for VQ design," in *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 1996, pp. 2048–2051.

[37] T. Berger, *Rate-distortion theory*. Englewood Cliffs, NJ, USA: Prentice Hall, 1971.

[38] M. Skoglund, "On channel-constrained vector quantization and index assignment for discrete memoryless channels," *IEEE Trans. on Information Theory*, vol. 45, no. 7, pp. 2615–2622, November 1999.

Paper C

Lattice-Based Source–Channel Coding in Wireless Sensor Networks

Johannes Karlsson and Mikael Skoglund

To appear in *Proceedings of IEEE International Conference on Communications*, June 2011.

Lattice-Based Source–Channel Coding in Wireless Sensor Networks

Johannes Karlsson and Mikael Skoglund

Abstract

We consider the problem of gathering measurements in a wireless sensor network consisting of a large number of sensor nodes. A practical joint source–channel coding scheme is proposed and evaluated. The scheme uses lattices to extend a previously proposed scheme to higher dimensions. The key idea is to use conventional point-to-point communication for a subset of the sensor nodes and side-information aware transmission for the remaining sensor nodes. The selection of sensors is based on their instantaneous channel quality. It is shown that by expanding from one to eight dimensions, a gain of about 1 *dB is achievable. The overall transmission delay of the scheme is still very low and it is therefore suitable to use in delay-sensitive applications.*

1 Introduction

We consider the important task of communicating correlated sensor measurements to a fusion center over noisy channels. We will assume that the sensor measurements are real and thus take values from an infinite alphabet. This is a lossy distributed source–channel coding problem. Lossy source coding with side-information was first studied by Wyner and Ziv in [1], where they characterize the rate–distortion region when side information is available at the decoder. This problem has been extended to other scenarios of lossy distributed source coding, such as multi-terminal source coding [2, 3] and the Chief Executive Officer (CEO) problem [4]. In multi-terminal source coding in general, the objective is to estimate an entire field at the fusion center, where each sensor can measure only a part of the field. Whereas in the CEO problem, all sensors have noisy observations of the same source and the objective is to estimate this single underlying

source at the fusion center. In the following we will consider the first of these problems, namely, to estimate an entire field.

The rate–distortion region of this problem is only known in a few special cases, such as the quadratic Gaussian two-encoder problem [5]. It is, in general, very difficult to find the rate-distortion region for larger networks. Nevertheless, there are numerous upper and lower bounds, see [2, 3, 5–7].

In this paper, we will assume that the communication channels are made orthogonal by either frequency- or time-division multiple access techniques. It was recently proved [8] that the separation of source and channel coding, which is optimal in point-to-point communication, also is optimal in this scenario. However, the focus of our work is on *delay-sensitive* applications, such as networked control systems. In these applications the measurements are used in active decisions and there is a penalty involved in receiving delayed measurements; the system might even become unstable in the worst case. The information theoretic arguments for optimal source and channel coding and the separation principle, are all based on block lengths that asymptotically tend to infinity. It is clear that long block lengths, which would lead to significant decoding delays, does not fulfill our low-delay constraints. Therefore, we propose the use of a joint source–channel code, operating on short blocks of only a few samples, which gives virtually no transmission delay at all.

Uncoded linear transmission has been shown to outperform schemes based on separate source and channel codes if specific conditions are fulfilled [9]. Because of this, linear transmission may seem to be a good candidate of a joint source–channel code for the problem at hand, see [10–12] for examples where this is used. However, since the channels are *orthogonal* and the measurements are *correlated* (as in [10–12]), a linear transmission scheme will only use a subspace of the channel; in the limit of high correlation, the subspace will be one-dimensional, which clearly is a waste of energy.

In [13], we considered a pair of sensors and proposed a design algorithm that jointly, in an iterative manner, optimized the operation at each sensor node as well as the decoder at the fusion center. This resulted in highly optimized systems that, in general, were nonlinear. The nonlinearities were such that all dimensions of the channel space were utilized. However, although straightforward in theory, the generalization to more than a couple of sensor nodes or vector measurements is infeasible due to high computational complexity. To reduce the complexity, we need some structure. A practical analog scheme that uses piecewise linear mappings was presented in [14]. The simple encoding and decoding structure made it possible to handle an arbitrary number of sensor nodes.

We generalize the ideas of piecewise linear mappings to higher dimensions by using modulo-lattice modulation (MLM). In addition, we also include fading channels in our model and present an idea for scheduling of

the sensor nodes. Lattices have been proposed and analyzed as a tool to do structured distributed source coding in WSNs before, see for example [15–18]. However, most results are theoretical and asymptotic in some sense. The idea of MLM originates from nested lattice codes in digital systems [17, 18] and was mentioned in [15] as a means of performing analog communication with side information. The idea is analyzed in [16] in a point-to-point scenario, where it is shown that MLM asymptotically reaches the Wyner–Ziv bound as the lattice dimension goes to infinity. Another motivation for MLM in situations where side information is available at the decoder can be found in [19]. In this case a scalar relay mapping was numerically optimized and the resulting mapping is very similar to a sawtooth curve, which is the one-dimensional equivalence of MLM. Motivated by low-delay constraints, we will look at finite-dimensional systems and especially investigate the gains in a practical system when the dimensions are increased. In what follows, we propose a practical and scalable lattice-based strategy for collecting sensor measurements in a WSN consisting of a few to hundreds of sensors.

2 Problem Formulation

We study a complete WSN with K sensor nodes as illustrated in Figure 1. Each sensor measures an N-dimensional quantity

$$\boldsymbol{Y}_i = \boldsymbol{X} + \boldsymbol{W}_i \in \mathbb{R}^N \qquad i = 1, \ldots, K, \tag{1}$$

where $\boldsymbol{X}$ and $\boldsymbol{W}_i$ are independent Gaussian random variables with covariance matrices $\sigma_X^2 \boldsymbol{I}$ and $\sigma_{W_i}^2 \boldsymbol{I}$, respectively. Hence, the samples are spatially correlated but temporally uncorrelated. We would like to emphasize that we are considering the multi-terminal source coding problem of estimating a random field, which is modeled by $\boldsymbol{Y}_i$, $i = 1, \ldots, K$. That is, the objective is to estimate *each* sensor's observation at the fusion center and we are *not* interested in $\boldsymbol{X}$, except to the extent it can be helpful in estimating $\boldsymbol{Y}_i$, $i = 1, \ldots, K$. Each sensor encodes its measurement independently in blocks of length N according to the function

$$\boldsymbol{s}_i(\boldsymbol{y}_i) : \mathbb{R}^N \mapsto \mathbb{R}^N \qquad i = 1, \ldots, K, \tag{2}$$

which is then transmitted to the fusion center over a Rayleigh block fading channel with additive white Gaussian noise (AWGN). The received value from each sensor is denoted by $\boldsymbol{R}_i$ and can be expressed as

$$\boldsymbol{R}_i = H_i \boldsymbol{s}_i(\boldsymbol{Y}_i) + \boldsymbol{V}_i \qquad i = 1, \ldots, K, \tag{3}$$

where H_i is the real-valued Rayleigh fading coefficient with $\mathbf{E}[|H_i|^2] = 1$ and $\boldsymbol{V}_i$ is Gaussian noise with covariance matrix $\sigma_V^2 \boldsymbol{I}$. At this point, we assume for simplicity that perfect channel-state information (CSI) is globally

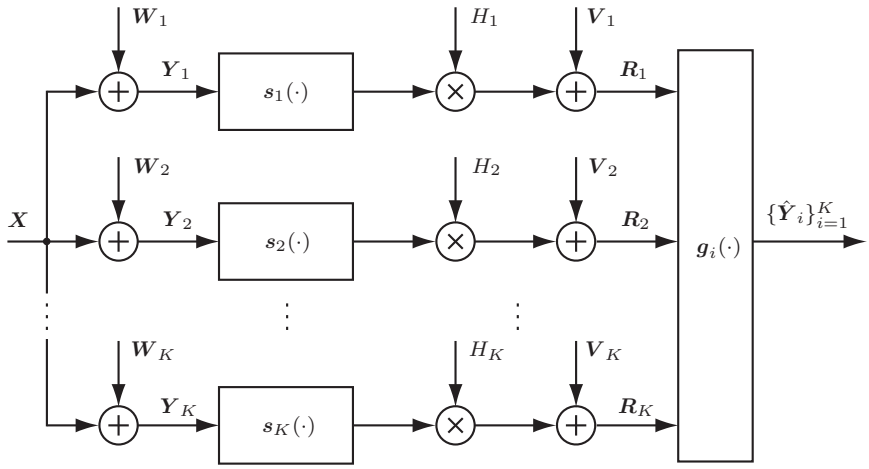


Figure 1: Overview of the WSN. K sensor nodes that should communicate their measurements to a fusion center.

available. The fusion center estimates each $\boldsymbol{Y}_i$ from all received values as,

$$\hat{\boldsymbol{Y}}_i = \boldsymbol{g}_i(\boldsymbol{R}_1, \boldsymbol{R}_2, \ldots, \boldsymbol{R}_K) \qquad i = 1, \ldots, K. \tag{4}$$

To have a low-delay system we want the block length N to be small. Typically, this will affect the performance and there is a tradeoff between delay and performance. As performance measure we use the mean squared error (MSE) per sample, defined as

$$\text{MSE} \triangleq \frac{1}{KN} \sum_{i=1}^{K} \mathbf{E}[\|\boldsymbol{Y}_i - \hat{\boldsymbol{Y}}_i\|^2]. \tag{5}$$

The objective is to minimize the MSE by a proper choice of encoding and decoding functions $\{\boldsymbol{s}_i, \boldsymbol{g}_i\}_{i=1}^K$, where $\boldsymbol{s}_i$ should satisfy an average power constraint

$$\frac{1}{N}\mathbf{E}[\|\boldsymbol{s}_i(\boldsymbol{y}_i)\|^2] \leq P \qquad i = 1, \ldots, K. \tag{6}$$

3 Proposed Scheme

Due to fading, the signal-to-noise ratio (SNR) of the different channels to the fusion center will be different. Some sensors will have a high SNR, which could correspond to the sensor being close to the fusion center, line-of-sight transmission or constructive multipath propagation; whereas others

have a low SNR, which could correspond to the sensors being further away, something blocking their transmission or destructive multipath propagation. Distributed source coding is mainly a way to reduce transmission rate or equivalently, in analog transmission as studied in this paper, a way to mitigate the effects of the channel noise. Because of this and since we are interested in each sensor's measurement, there is little gain in using distributed source coding for the sensors that already have a high SNR. We therefore propose the following scheduling and transmission strategy:

1. Transmit the measurements from the M sensors having highest SNR with conventional point-to-point source–channel coding techniques. In this scenario we use linear transmission.
2. Estimate $\boldsymbol{X}$ and $\boldsymbol{Y}_i$, $i = 1, \ldots, M$, from these transmissions.
3. Transmit the measurements from the remaining sensors using MLM, assuming that the estimate of $\boldsymbol{X}$ is available as side information at the receiver.

3.1 Phase 1: Linear Transmission

Without loss of generality, let the sensors be in such way that linear transmission is used for the M first sensor nodes, that is,

$$\boldsymbol{s}_i(\boldsymbol{y}_i) = \sqrt{\frac{P}{\sigma_X^2 + \sigma_{W_i}^2}}\boldsymbol{y}_i \qquad i = 1, \ldots, M. \tag{7}$$

The optimal receiver given the M linear transmissions is obtained by solving the Wiener–Hopf equations [20]. Let the elements of a vector $\boldsymbol{A}_i$ be denoted by A_{ij}, $j = 1, \ldots, N$ (e.g., $\boldsymbol{R}_i = [R_{i1}, R_{i2}, \ldots, R_{iN}]^T$). Each component of Y_i can now be estimated by

$$\hat{Y}_{ij} = \boldsymbol{t}^{(i)^T}\boldsymbol{R}^{(j)} \qquad i = 1, \ldots, M, j = 1, \ldots, N, \tag{8}$$

where

$$\boldsymbol{R}^{(j)} = [R_{1j}, R_{2j}, \ldots, R_{Kj}] \qquad j = 1, \ldots, N \tag{9}$$

$$\boldsymbol{t}^{(i)} = \left(\mathbf{E}[\boldsymbol{R}^{(j)}\boldsymbol{R}^{(j)^T}]\right)^{-1}\mathbf{E}[Y_{ij}\boldsymbol{R}^{(j)}] \qquad i = 1, \ldots, M. \tag{10}$$

In a similar manner, the optimal estimate of $\boldsymbol{X}$ is given by

$$\hat{X}_j = \boldsymbol{t}^{(0)^T}\boldsymbol{R}^{(j)} \qquad j = 1, \ldots, N, \tag{11}$$

where

$$\boldsymbol{t}^{(0)} = \left(\mathbf{E}[\boldsymbol{R}^{(j)}\boldsymbol{R}^{(j)^T}]\right)^{-1}\mathbf{E}[X_j\boldsymbol{R}^{(j)}]. \tag{12}$$

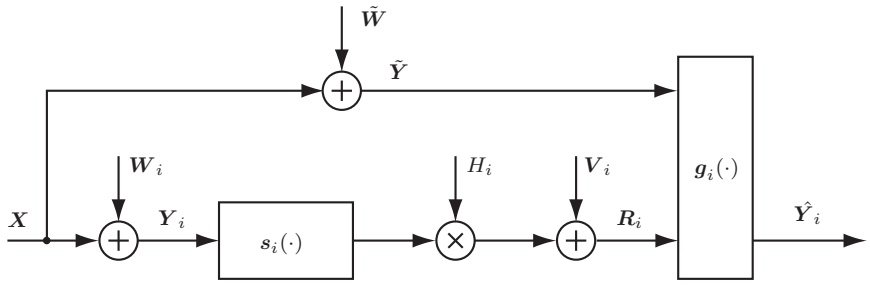


Figure 2: Transmission from one isolated sensor node using MLM with side information at the receiver.

3.2 Phase 2: Modulo-Lattice Modulation

We need some basic terminology for lattices before we can proceed and look at the modulo-lattice modulation. An N-dimensional lattice Λ is defined by the generator matrix $G \in \mathbb{R}^{N\times N}$. A point $\boldsymbol{l} \in \mathbb{R}^N$ belongs to the lattice if and only if it can be written as $\boldsymbol{l} = G\boldsymbol{i}$, where $\boldsymbol{i} \in \mathbb{Z}^N$ and $\mathbb{Z} = \{0, \pm 1, \pm 2, \ldots\}$. We define the nearest neighbor quantizer of a lattice Λ by

$$Q_{\Lambda_\Delta}(\boldsymbol{x}) \triangleq \arg\min_{\boldsymbol{l}\in\Lambda_\Delta} \|\boldsymbol{x} - \boldsymbol{l}\|, \tag{13}$$

where Δ is used to explicitly specify the minimum distance between lattice points (which is the same as two times the packing radius of the lattice). For a particular lattice Λ, any Δ can be achieved by a proper scaling of G. For the lattices considered in this paper, there exist highly efficient algorithms that perform the operation in (13) without the need of an exhaustive search. The modulo-lattice operation that will be used in the rest of the paper is defined by

$$\boldsymbol{x} \bmod \Lambda_\Delta \triangleq \boldsymbol{x} - Q_{\Lambda_\Delta}(\boldsymbol{x}). \tag{14}$$

A comprehensive study of lattices and lattice quantization can be found in [21]. See also [16], where the MLM scheme is presented and analyzed in detail.

We will now present how we will use MLM in our scenario. Given the M linear transmissions, we will isolate each of the remaining $K-M$ sensors and treat them individually as shown in Figure 2. The motivation for the individual treatment of the remaining sensors is the fact that $\boldsymbol{Y}_i$ and $\boldsymbol{Y}_j$ are conditionally independent given $\boldsymbol{X}$ for all $j \neq i$. The side-information from the linear transmissions is modeled by $\tilde{\boldsymbol{Y}} = \boldsymbol{X} + \tilde{\boldsymbol{W}}$, where $\tilde{\boldsymbol{W}}$ is Gaussian noise with covariance matrix $\sigma^2_{\tilde{W}}\boldsymbol{I}$[1]. We use the correlation coefficients ρ_{XY}

[1] $\tilde{\boldsymbol{Y}}$ can be obtained by a proper scaling of $\hat{\boldsymbol{X}}$ in (11).

and $\rho_{X\tilde{Y}}$ to quantify the quality of the side-information,

$$\rho_{XY} \triangleq \frac{\mathbf{E}[X_j Y_{ij}]}{\sigma_X \sigma_{Y_i}}, \rho_{X\tilde{Y}} \triangleq \frac{\mathbf{E}[X_j \tilde{Y}_j]}{\sigma_X \sigma_{\tilde{Y}}}. \tag{15}$$

Since side-information is available at the fusion center, we can use the following nonreversible modulo-operation at the source nodes

$$\boldsymbol{s}_i(\boldsymbol{y}_i) = \beta_i(\alpha_i \boldsymbol{y}_i \bmod \Lambda_{\Delta_i}) \qquad i = M+1, \ldots, K. \tag{16}$$

There are three scaling parameters involved — α_i, β_i, and Δ_i. α_i is chosen such that the variance of each component of $\alpha_i \boldsymbol{Y}_i$ equals unity. In a similar way, β_i is used to make sure that the power constraint in (6) is fulfilled. The parameter Δ_i specifies the minimum distance of lattice points and determines the amount of information that is removed prior to transmission. If Δ_i is large no information is removed and the sensor operates in a linear mode. On the other hand, if Δ_i is small, more information is removed and the fusion center would have to rely on the side-information to estimate the removed part and be able to decode.

The decoder is divided into three steps. First, the output of the sensor node, $\boldsymbol{s}_i(\boldsymbol{y}_i)$, is estimated from $\boldsymbol{R}_i$ by using the linear minimum MSE (MMSE) estimator,

$$\hat{\boldsymbol{s}}_{i|R} = \frac{Ph_i}{Ph_i^2 + \sigma_{V_i}^2} \boldsymbol{R}_i \qquad i = M+1, \ldots, K. \tag{17}$$

Next, the part that is removed during the modulo operation is estimated from $\hat{\boldsymbol{X}}$ and $\hat{\boldsymbol{s}}_{i|R}$,

$$\hat{Q_{\Lambda i}} = Q_{\Lambda_{\Delta_i}}(\alpha_i \hat{\boldsymbol{X}} - \frac{\hat{\boldsymbol{s}}_{i|R}}{\beta_i}) \qquad i = M+1, \ldots, K. \tag{18}$$

Finally, $\boldsymbol{Y}_i$ is estimated as

$$\hat{\boldsymbol{Y}}_i = \frac{1}{\alpha_i}(\hat{Q_{\Lambda i}} + \frac{\hat{\boldsymbol{s}}_{i|R}}{\beta_i}) \qquad i = M+1, \ldots, K. \tag{19}$$

Provided that the side-information is sufficiently good such that $\hat{Q_{\Lambda i}} = Q_{\Lambda_{\Delta_i}}(\alpha_i \boldsymbol{Y}_i)$ with high probability, decreasing Δ_i would allow the sensor to increase β_i without violating the power constraint. This can be seen as a virtual power gain and decreases the MSE. However, if Δ_i is decreased below some threshold such that decoding errors become dominant in (18), the MSE will increase. The optimal Δ_i will depend both on ρ_{XY} and $\rho_{X\tilde{Y}}$ as well as the instantaneous SNR, $Ph_i^2/\sigma_{V_i}^2$. We have performed a numerical grid search to find the optimal Δ_i for a large range of these parameters and

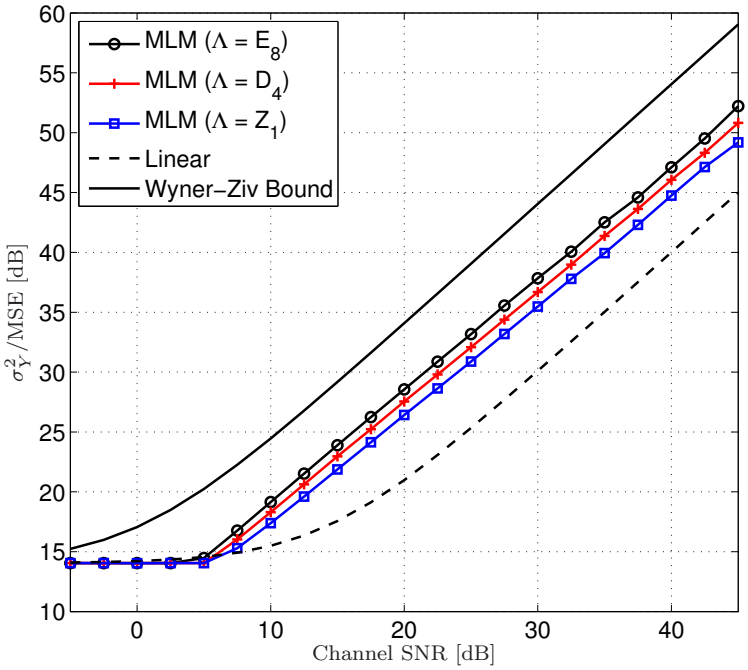


Figure 3: Numerical results of single-link MLM using the lattices E_8, D_4 and Z_1. Linear transmission and the Wyner–Ziv bound are shown for comparison. $\rho_{XY} = \rho_{X\hat{Y}} = 0.99$

the lattices E_8, D_4, Z_1, where the subindices in this case denote the lattice dimension. The resulting performance for $\rho_{XY} = \rho_{X\tilde{Y}} = 0.99$ can be seen in Figure 3, where we also compare the results to linear transmission and the Wyner–Ziv bound. The lattices we use — Z_1, D_4, and E_8 — are well known for their excellent properties. All of them are the densest lattice packings in their respective dimension and it is "a reasonable guess" [21] that they are also the best uniform quantizers among lattices. See [21] for details about their definitions and properties.

By looking at Figure 3 we can see that linear transmission is slightly better than MLM for SNR points less than 7–8 dB. Because of this we make a small modification to the selection of sensors that use linear transmission and include also those sensors with low SNR, where linear transmission is superior. As the SNR increases we see that the curves are more or less parallel. For the specified correlation, MLM with the lattice E_8 is 7–8 dB better than linear transmission. The gain of increasing the dimensions from 1 to 4 and from 4 to 8 is roughly 1 dB. The gap to the Wyner–Ziv bound is 6–7 dB, however, it is worth to stress that this bound assumes infinite dimensions in both the source and channel coding. In [16] it is in fact shown that MLM can reach the Wyner–Ziv bound as the lattice dimension goes to infinity.

4 Numerical Results

To evaluate the proposed method we will consider two scenarios — a WSN with 20 sensors and a WSN with 100 sensors. For each of these scenarios we let $\sigma^2_{W_i} = \sigma^2_W$ and simulate the system for two different correlation coefficients, namely $\rho_{XY} = 0.99$ and $\rho_{XY} = 0.995$. For the Rayleigh fading, we have assumed that the channel coefficients stay constant for 100 channel uses. The parameter M, which determines the number of sensors using linear transmission, has been optimized numerically for each SNR point. The results of the simulations for the two scenarios are presented in Figure 4 and Figure 5. It is evident that our proposed method clearly outperforms transmission using linear modulation. As soon as the SNR exceeds a certain threshold, the gains are significant. This is despite the fact that the optimal receiver is used for the linear scheme whereas for the MLM scheme we have resorted to a suboptimal ad-hoc receiver. The limitation of the linear system is on the transmitting side. In the limit of a correlation coefficient $\rho_{XY} = 1$, for each sample, linear transmission only uses a one-dimensional subspace of the K-dimensional channel space. This is the explanation why the linear system is not able to take advantage of the increased correlation as the SNR increases; the curves for $\rho_{XY} = 0.99$ and $\rho_{XY} = 0.995$ asymptotically converge to the curve $\rho_{XY} = 0$. This is in contrast to the proposed system with MLM, where the advantage of a higher correlation is maintained as

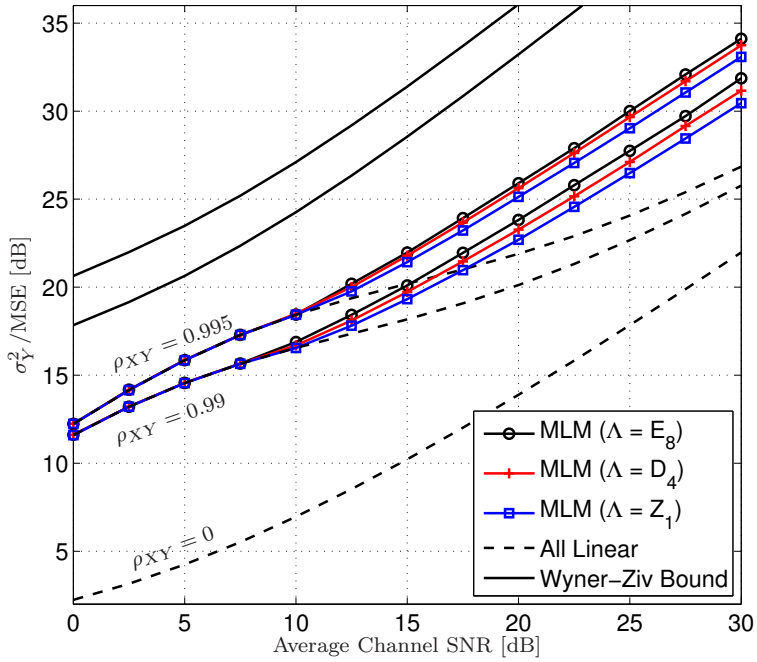


Figure 4: WSN consisting of 20 sensor nodes. MLM in comparison to Linear transmission for $\rho_{XY} = 0.99$ and $\rho_{XY} = 0.995$.

the SNR is increased. This is achieved by a decrease of Δ which makes sure that all dimensions of the channel space are used.

The Wyner–Ziv bound is obtained from Theorem 3 in [5], where the minimum MSE given a sum-rate constraint is stated. The bound is used by taking the sum of each link's ergodic channel capacity as the sum rate. In general, this bound is too optimistic and is not achievable with short block lengths since it relies on infinite block lengths in several senses. For example, there is no outage since we use the ergodic channel capacity, the effects of the channel noise is cancelled out by an optimal channel code, and the correlation behaves exactly as the correlation coefficient predicts.

Another observation is that the threshold where MLM performs better than linear transmission is shifted towards a lower SNR for the scenario with 100 sensors. This is explained by looking at Figure 6, where the number of linear sensors are plotted as a function of the SNR. At an SNR of 7.5 dB all sensors use linear transmission in the WSN with 20 sensors, while around

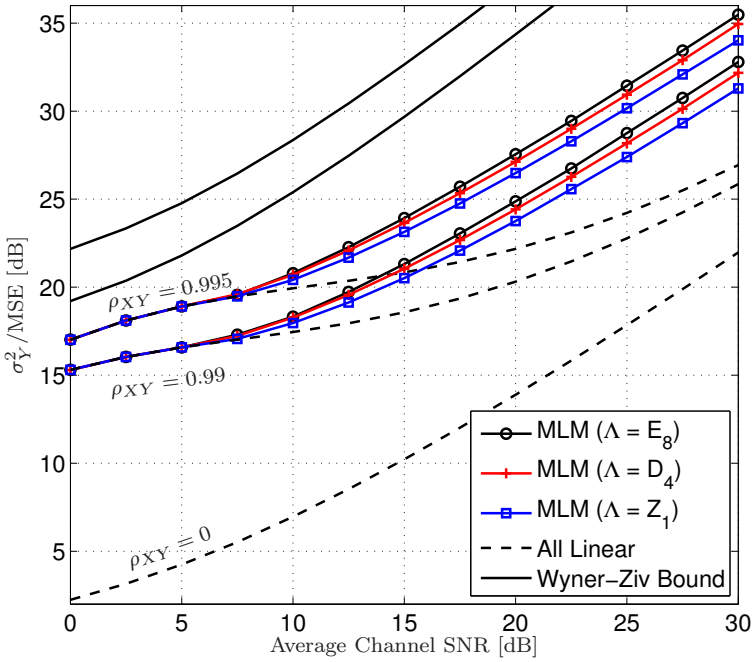


Figure 5: WSN consisting of 100 sensor nodes. MLM in comparison to Linear transmission for $\rho_{XY} = 0.99$ and $\rho_{XY} = 0.995$.

40 sensors use MLM in the WSN consisting of 100 sensor nodes. To put it in another way, as the number of sensor nodes increases, a smaller amount of sensors are needed to provide sufficiently good side information which is necessary for MLM to work.

5 Conclusions

We have proposed a joint source–channel coding scheme for WSNs based on MLM. The scheme is practical and has a low complexity that makes it suitable for WSNs consisting of a large number of sensors nodes. The use of very short block lengths is favorable in delay-sensitive applications such as control systems. We have investigated the gains of increasing the dimensions and shown that an 8-dimensional system performs about 1 dB better than a scalar system; linear transmission is clearly outperformed in all cases. The assumption of globally available CSI can easily be relaxed

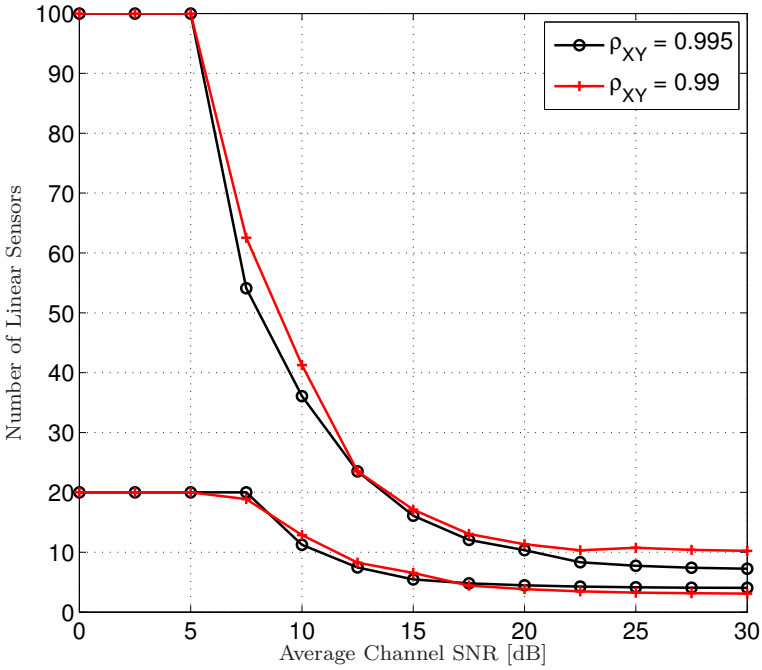


Figure 6: Number of sensors that use linear transmission as a function of average channel SNR for MLM with E_8. $\rho_{XY} = 0.99$ and $\rho_{XY} = 0.995$.

to a more realistic assumption of local CSI. The fusion center could then broadcast a threshold that determines if MLM or linear transmission should be used.

References

[1] A. D. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. on Information Theory*, vol. 22, no. 1, pp. 1–10, January 1976.

[2] T. Berger, "Multiterminal source coding," in *The information theory approach to communications*, G. Longo, Ed. Springer–Verlag, 1977.

[3] S. Tung, "Multiterminal source coding," Ph.D. dissertation, Cornell University, Ithaca, NY, 1978.

[4] T. Berger, Z. Zhang, and H. Viswanathan, "The CEO problem [multiterminal

source coding]," *IEEE Trans. on Information Theory*, vol. 42, no. 3, pp. 887–902, May 1996.

[5] A. B. Wagner, S. Tavildar, and P. Viswanath, "Rate region of the quadratic Gaussian two-encoder source-coding problem," *IEEE Trans. on Information Theory*, vol. 54, no. 5, pp. 1938–1961, May 2009.

[6] A. B. Wagner and V. Anantharam, "An improved outer bound for multiterminal source coding," *IEEE Trans. on Information Theory*, vol. 54, no. 5, pp. 1919–1937, May 2008.

[7] Y. Oohama, "Distributed source coding of correlated Gaussian observations," in *Proceedings IEEE Int. Symp. Information Theory and its Applications*, December 2008.

[8] J. J. Xiao and Z. Q. Luo, "Multiterminal source-channel communication over an orthogonal multiple-access channel," *IEEE Trans. on Information Theory*, vol. 53, no. 9, pp. 3255–3264, September 2007.

[9] M. Gastpar, B. Rimoldi, and M. Vetterli, "To code, or not to code: lossy source-channel communication revisited," *IEEE Trans. on Information Theory*, vol. 49, no. 5, pp. 1147–1158, May 2003.

[10] S. Cui, J. Xiao, A. Goldsmith, Z. Luo, and H. V. Poor, "Estimation diversity and energy efficiency in distributed sensing," *IEEE Trans. on Signal Processing*, vol. 55, no. 9, pp. 4683–4695, 2007.

[11] I. Bahceci and A. K. Khandani, "Linear estimation of correlated data in wireless sensor networks with optimum power allocation and analog modulation," *IEEE Trans. on Communications*, vol. 56, no. 7, pp. 1146–1156, 2008.

[12] J. Fang and H. Li, "Power constrained distributed estimation with correlated sensor data," *IEEE Trans. on Signal Processing*, vol. 57, no. 8, pp. 3292–3297, 2009.

[13] N. Wernersson, J. Karlsson, and M. Skoglund, "Distributed quantization over noisy channels," *IEEE Trans. on Communications*, vol. 57, no. 6, pp. 1693–1700, June 2009.

[14] N. Wernersson and M. Skoglund, "Nonlinear coding and estimation for correlated data in wireless sensor networks," *IEEE Trans. on Communications*, vol. 57, no. 10, pp. 2932–2939, October 2009.

[15] Z. Reznic, M. Feder, and R. Zamir, "Distortions bounds for broadcasting with bandwidth expansion," *IEEE Trans. on Information Theory*, vol. 52, no. 8, pp. 3778–3788, August 2006.

[16] Y. Kochman and R. Zamir, "Joint Wyner–Ziv/dirty-paper coding by modulo-lattice modulation," *IEEE Trans. on Information Theory*, vol. 55, no. 11, pp. 4878–4889, November 2009.

[17] R. Zamir, S. S. (Shitz), and U. Erez, "Nested linear/lattice codes for structured multiterminal binning," *IEEE Trans. on Information Theory*, vol. 48, no. 6, June 2002.

[18] S. D. Servetto, "Lattice quantization with side information: codes, asymptotics and applications in sensor networks," *IEEE Trans. on Information Theory*, vol. 53, pp. 714–731, February 2007.

[19] J. Karlsson and M. Skoglund, "Optimized low-delay source–channel–relay mappings," *IEEE Trans. on Communications*, no. 5, pp. 1397–1404, May 2010.

[20] S. M. Kay, *Fundamentals of statistical signal processing: estimation theory.* Upper Saddle River, NJ, USA: Prentice–Hall, Inc., 1993.

[21] J. H. Conway and N. J. A. Sloane, *Sphere packings, lattices and groups.* New York, USA: Springer-Verlag, 1988.

Paper D

Analog Distributed Source–Channel Coding Using Sinusoids

Johannes Karlsson and Mikael Skoglund

Published in *Proceedings of International Symposium on Wireless Communication Systems* (invited paper), September 2009.

Analog Distributed Source–Channel Coding Using Sinusoids

Johannes Karlsson and Mikael Skoglund

Abstract

In a wireless sensor network, it is likely that the measurements of the sensors are correlated. Distributed source coding can be used to reduce transmission rate or mitigate the effects of the channel noise in the case of analog transmission. In this paper, we propose a novel scheme for implementing distributed source–channel coding based on analog mappings. We assume that an analog source is to be transmitted to a receiver that has access to correlated side information, as in the Wyner–Ziv problem. From the Cramér–Rao lower bound, we observe general properties of analog distributed source–channel mappings. It is especially clear how the stretch factor influences the performance. From this observation we propose two different mappings based on sinusoidal waveforms. The proposed transmission scheme is numerically evaluated and shown to perform well, particularly in the low-SNR regime. Furthermore, it requires no encoding or decoding delay, making it suitable for delay-critical applications in wireless sensor networks.

1 Introduction

Over the last decade, wireless sensor networks (WSNs) have received considerable attention. A WSN should consist of small, cheap, and energy-efficient devices which measure relevant quantities. Due to high spatial density of the sensors, the measurements are likely to be correlated. One important task is to find an efficient way of communicating all sensor measurements to a fusion center.

Ever since the work by Wyner and Ziv [1], different ways of implementing lossy distributed source coding have been proposed. The most common approach is to do quantization followed by Slepian–Wolf coding [2]. The Slepian–Wolf code is often implemented using long block codes which introduces significant delays in the decoding process. In situations, such as

control applications, where a low delay is of interest, there is much to gain by using a joint source–channel code, where the operations at the sensor node are merged to one single operation — a mapping from the source space directly to the channel space. In what follows, we propose a novel scheme for distributed source–channel coding based on analog mappings, which results in very low transmission delays.

2 Problem Formulation

Consider a wireless sensor network with K sensors, where the measurements of each sensor are correlated according to

$$X_i = Y + N_i \qquad i = 1, \ldots, K, \tag{1}$$

where Y and N_i are independent Gaussian random variables with variances σ_Y and σ_{N_i}, respectively. The objective is to transmit each sensors' individual reading (i.e., X_i) to a fusion center. The variables Y and N_i are only used to describe the correlation structure. Due to the physical placement of the different sensors, the signal-to-noise ratio (SNR) of their channels to the fusion center will be different. Some sensors will be close to the fusion center and hence have a high SNR, whereas others are further away or for some other reason have a low SNR. Distributed source coding is mainly a way to reduce transmission rate or equivalently, in analog transmission as studied in this paper, a way to mitigate the effects of the channel noise. Because of this there is not much gain in using distributed source coding for the sensors that already have a high SNR. We therefore propose the following transmission strategy: 1. Transmit the measurements from the sensors with an SNR higher than some threshold with conventional source–channel coding techniques. 2. Estimate Y from these measurements. 3. Transmit the measurements from the remaining sensors using distributed source–channel coding, assuming that the estimate of Y is available as side information at the receiver.

In this paper, we will only consider the third step and assume that Y has been perfectly estimated[1] and is available at the receiver. From now on we study the transmission and estimation of the measurement of the ith sensor. To simplify the notation we drop the index specifying the sensor and consequently have a situation where the sensor measures

$$X = Y + N. \tag{2}$$

[1]This is a slight simplification since the estimate of Y will not be perfect. However, this could be compensated by letting σ_N be slightly larger than σ_{N_i} in the following discussion.

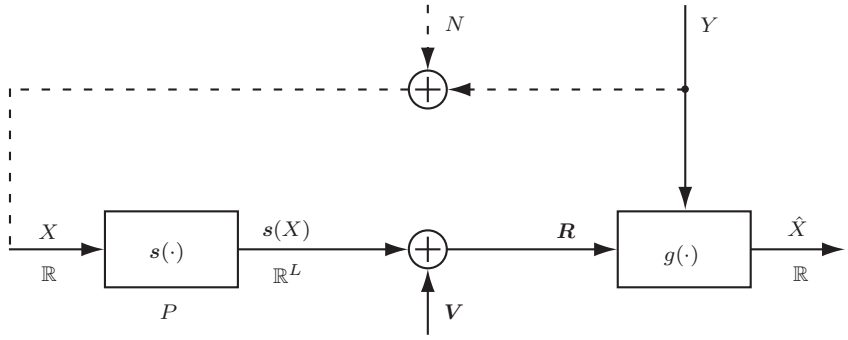


Figure 1: Analog transmission from a sensor node with side information at the receiver. The dashed line shows the structure of the correlation.

The correlation between X and Y is specified by the correlation coefficient ρ,

$$\rho = \frac{\mathbf{E}[XY]}{\sigma_X \sigma_Y} = \frac{\sigma_Y}{\sigma_X}. \tag{3}$$

The channel from the sensor to the fusion center is modeled as a memoryless additive white Gaussian noise (AWGN) channel which can be used L times for each measurement. The sensor encodes X using a mapping $\boldsymbol{s}(X) : \mathbb{R} \mapsto \mathbb{R}^L$. The mapping $\boldsymbol{s}(X)$ should be chosen such that the average power constraint

$$\mathbf{E}[\|\boldsymbol{s}(X)\|^2] \leq LP \tag{4}$$

is satisfied. The fusion center receives

$$\boldsymbol{R} = \boldsymbol{s}(X) + \boldsymbol{V}, \tag{5}$$

where $\boldsymbol{V}$ is zero-mean Gaussian noise with covariance matrix $\mathbf{E}[\boldsymbol{V}^T\boldsymbol{V}] = \sigma_V^2 \boldsymbol{I}$. X is finally estimated from $\boldsymbol{R}$ and the side information Y by a mapping $\hat{X} = g(\boldsymbol{R}, Y) : \mathbb{R}^L \times \mathbb{R} \mapsto \mathbb{R}$. The system is shown in Figure 1. As performance measure we use the mean squared error

$$D = \mathbf{E}[(X - \hat{X})^2]. \tag{6}$$

The objective now is to find a mapping $\boldsymbol{s}(X)$, satisfying (4), and its corresponding receiver mapping, $g(\boldsymbol{R}, Y)$, such that the distortion D is minimized. In the following sections, we will discuss general properties of such mappings and evaluate some mappings that are based on sinusoidal waveforms.

3 Analysis

The Cramér–Rao lower bound (CRLB) [3] gives a lower bound on the distortion for any unbiased receiver mapping. Evaluated for our problem, the CRLB states that (see Appendix A)

$$D \geq \frac{1}{\frac{1}{\sigma_V^2}\mathbf{E}[\|\boldsymbol{s}'(X)\|^2] + \frac{1}{\sigma_N^2}}, \tag{7}$$

where

$$\|\boldsymbol{s}'(X)\| = \sqrt{\sum_{i=1}^{L} \left\{\frac{ds_i(X)}{dX}\right\}^2} \tag{8}$$

is the stretch factor [4]. From (7), it can be seen how the stretch factor determines how much the channel noise is attenuated; the stretch factor should be as high as possible to limit the effect of the channel noise. One obvious way to accommodate this is by increasing the transmit power but this would violate the power constraint in (4). Another option is to let $\boldsymbol{s}(X)$ be a mapping that twists and bends as much as possible, which means that its derivatives and, consequently, also the stretch factor are high. There is a limit to how much the curve can be stretched by twisting and bending. At some point, different folds of the curve will come too close to each other such that a small channel noise may cause a large decoding error. In conventional source–channel coding, the curve should be stretched as much as possible, keeping different folds of the curve at a maximum distance to avoid large decoding errors [4, 5]. In the case of distributed source coding, the side information adds an extra dimension making it possible to have mappings that repeat themselves.

3.1 Optimal Receiver

The receiver should form an estimate of x based on the received vector $\boldsymbol{r}$ and the side information y. It is well-known that the optimal estimate, in minimum MSE sense, is given by the conditional expected value

$$\hat{x} = \mathbf{E}[X|\boldsymbol{r}, y]. \tag{9}$$

It should be noted that the conditional expected value in general is biased. Nevertheless, the CRLB can still give useful insights into the design of good encoder mappings.

4 Proposed Schemes

In this section, we propose two analog distributed source–channel mappings for the case $L = 2$. Our proposed mappings are based on sinusoidal wave-

forms that thanks to their periodic nature stretch the curve by reusing output symbols. Sinusoids have previously been used for bandwidth expansion of a uniform source with no side information in [7]. In that case the periodic nature of the sinusoids caused some problems and the system had to be designed with a safe margin such that the output symbols does not repeat. However, in our scenario the side information adds an extra dimension making the reuse of output symbols possible.

Mapping 1 is given by

$$\boldsymbol{s}(x) = A \begin{bmatrix} \cos(kx) \\ \sin(kx) \end{bmatrix} \tag{10}$$

$$\|\boldsymbol{s}'(x)\| = Ak \tag{11}$$

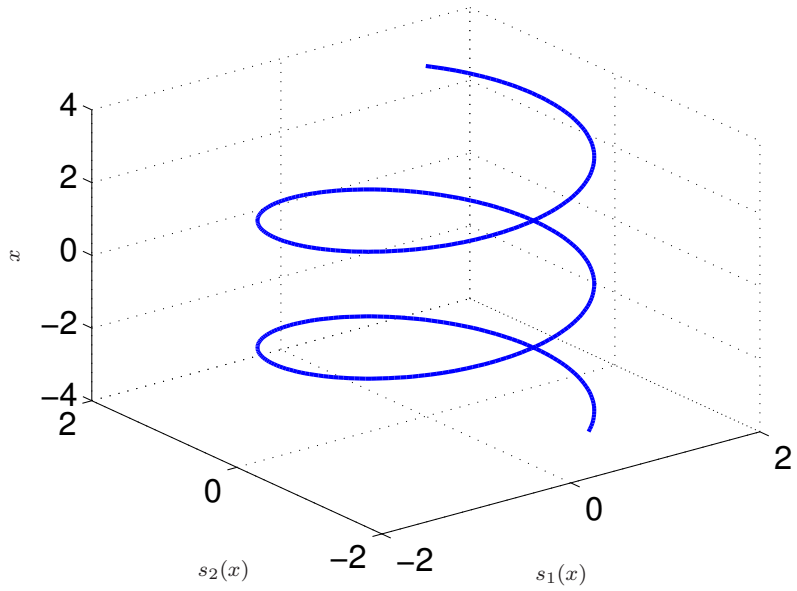
where we have also included its stretch factor. The mapping has two parameters, k and A, where k is used to control the periodicity and A is determined by the power constraint in (4). Two examples of this mapping are shown in Figure 2(a) and (b). The mapping in (b), with $k = 5.5$, would be preferable to the one in (a), with $k = 1.8$, due to its higher stretch factor. A direct implication of the higher value of k is that the different turns of the spiral are packed more closely, meaning that the receiver would have to put more trust in the side information. If the correlation is high, this would work well but if the correlation is low, there is a risk for large decoding errors and the mapping in (a) would be more suitable to use.

During simulations we noted that Mapping 1 suffered from performance saturation as the SNR increased. This is because the distortion in the case of a high SNR is dominated by errors that occur due to errors in the side information. That is, the transmitted value is estimated as coming from the wrong turn on the spiral. To avoid this behavior, we propose a second mapping, which is a generalization of Mapping 1 with a radius that varies with x. **Mapping 2** is given by

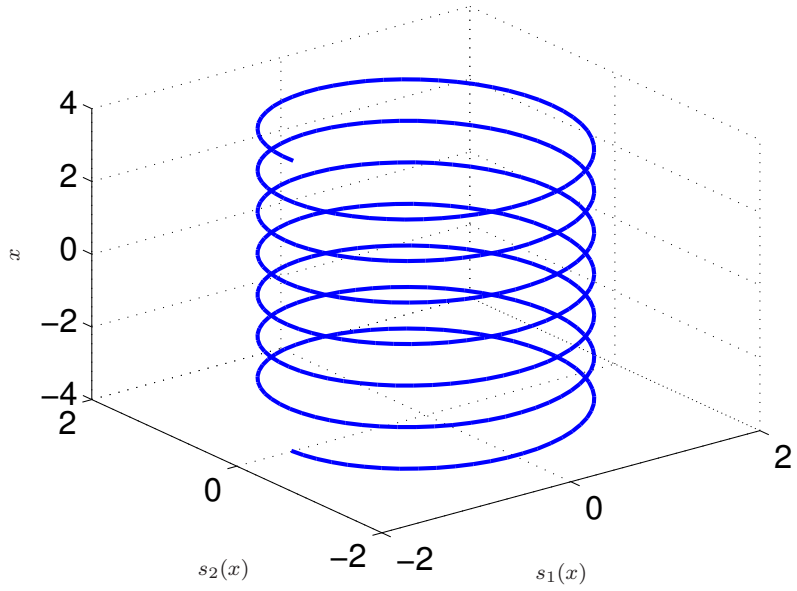
$$\boldsymbol{s}(x) = A(1 + r\cos(mkx)) \begin{bmatrix} \cos(kx) \\ \sin(kx) \end{bmatrix} \tag{12}$$

$$\|\boldsymbol{s}'(x)\| = Ak\sqrt{(1 + r\cos(mkx))^2 + (rm\sin(mkx))^2}. \tag{13}$$

As can be seen, the mapping has two additional parameters, r and m, where r determines how much the radius should vary from unity and m determines the frequency in relation to k at which the radius vary. By letting m be a noninteger, the mapping is repeated less frequently which gives better performance. An example of this mapping can be seen in Figure 3. In this figure $m = 4.5$, meaning that the mapping is repeated every second turn (this is due to the decimal part being .5). One drawback with Mapping 2 is the extra number of parameters, which makes it more complicated to find the optimal combination of parameters.



(a)



(b)

Figure 2: (a) Mapping 1: $k = 1.8$ (b) Mapping 1: $k = 5.5$

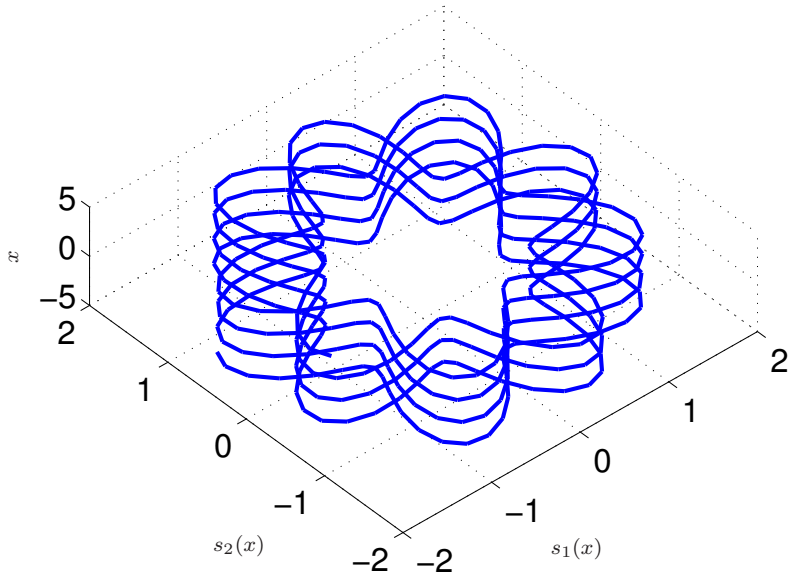


Figure 3: Mapping 2: $k = 5.6$, $m = 4.5$, $r = 0.35$

5 Numerical Results

The first step in evaluating our proposed mappings is to determine the optimal parameters to use. For Mapping 1 there is only one parameter to determine and we have used a grid search to find good values of k (See Table 1 in Appendix B). For Mapping 2 it is a bit more cumbersome to find the optimal parameters since the search space now is three dimensional. In this case we have constrained the search to the sets $m \in \{2.5, 3.5, \ldots, 16.5\}$ and $r \in \{0.2, 0.25, \ldots, 0.45\}$ (See Table 2 in Appendix Appendix B). We consider two different correlation coefficients in our simulations — $\rho = 0.9$ and $\rho = 0.99$. In the case of Mapping 2, we observe that the parameters m and r seem to only be dependent on the channel SNR and independent of ρ. The mappings in Figure 4(a) and Figure 4(b) are optimized for an SNR of 5 dB and 15 dB, respectively. We assume that the receiver knows the SNR and therefore adapts to the current channel conditions.

The mappings are evaluated against the 1:2-bandwidth-expansion mappings in [6] and also a linear scheme where

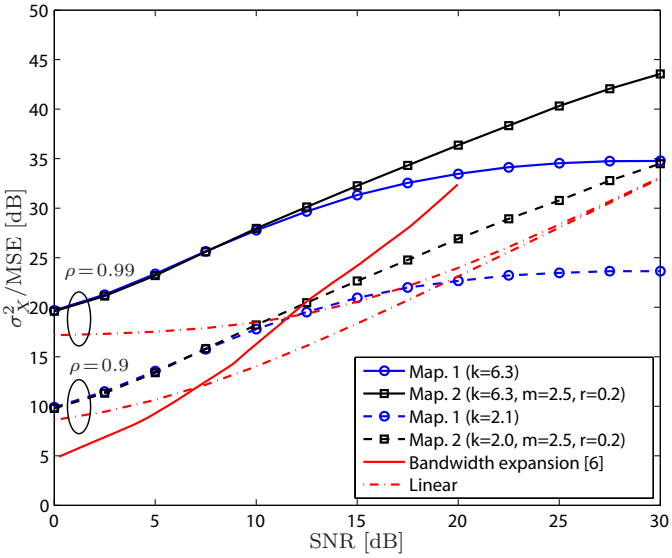
$$\boldsymbol{s}(x) = A \begin{bmatrix} x \\ x \end{bmatrix}. \tag{14}$$

The mappings in [6] does not utilize any side information but on the other hand they have a higher degree of freedom and have been optimized for each channel-SNR point. They therefore serve as a reference and are used to show the performance gains that come from using side information at the receiver. The linear system uses the optimal receiver, which makes use of the side information, and serves as a reference for the gains that come from a better use and reuse of the available output symbols.

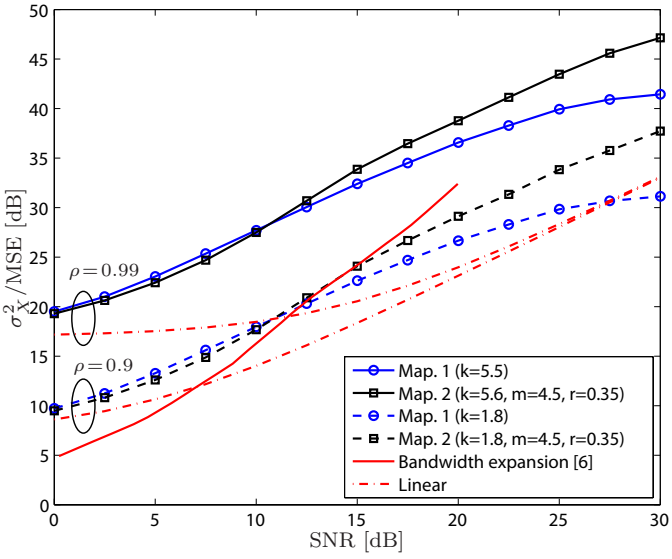
It is expected that the performance improves as the correlation coefficient increases. Interestingly, the proposed mappings manage to keep an approximately constant distance between the curves at $\rho = 0.9$ and $\rho = 0.99$; whereas for the linear system, the advantage of a higher correlation coefficient disappears as the SNR increases. We can also clearly see that Mapping 1 saturates earlier than Mapping 2 when the SNR increases.

6 Conclusions

We have proposed the use of analog source–channel mappings, based on sinusoidal waveforms, for implementing distributed source coding. The mappings have been numerically evaluated and are shown to perform well, especially in the case of high correlation and low channel SNR. The proposed transmission scheme requires no encoding or decoding delay, making it suitable for delay-critical applications. Possible directions for further work is to increase L and also investigate other mappings that can be used.



(a)



(b)

Figure 4: Performance evaluation. The mappings are optimized for (a) SNR=5 dB and (b) SNR=15 dB.

Appendix A Cramér–Rao Lower Bound

The Cramér–Rao Lower Bound [3] is given by

$$D \geq \left\{ -\mathbf{E}\left[\frac{\partial^2 \ln p(\boldsymbol{R}, Y, X)}{\partial X^2} \right] \right\}^{-1}$$
$$= \left\{ -\mathbf{E}\left[\frac{\partial^2 \ln p(\boldsymbol{R}|X)}{\partial X^2} + \frac{\partial^2 \ln p(Y|X)}{\partial X^2} + \frac{\partial^2 \ln p(X)}{\partial X^2} \right] \right\}^{-1}, \qquad (15)$$

where the equality follows from the factorization of $p(\boldsymbol{R}, Y, X)$, using the fact that $\boldsymbol{R}$ and Y are conditionally independent given X. The probability densities are given by

$$p(\boldsymbol{R}|X) = \frac{1}{(2\pi)^{L/2}\sigma_V^L} \exp\left\{ -\frac{1}{2\sigma_V^2} \sum_{i=1}^{L} [s_i(X) - R_i]^2 \right\} \qquad (16)$$

$$p(Y|X) = \frac{\sigma_X}{\sqrt{2\pi}\sigma_N\sigma_Y} \exp\left\{ -\frac{\sigma_X^2}{2\sigma_N^2\sigma_Y^2} \left[Y - \frac{\sigma_Y^2}{\sigma_X^2} X \right]^2 \right\} \qquad (17)$$

$$p(X) = \frac{1}{(2\pi)^{1/2}\sigma_X} \exp\left\{ -\frac{1}{2\sigma_X^2} X^2 \right\}. \qquad (18)$$

By taking the logarithm and thereafter deriving the second derivative with respect to X, (16)–(18) are transformed to

$$\frac{\partial^2 \ln p(\boldsymbol{R}|X)}{\partial X^2} = -\frac{1}{\sigma_V^2} \sum_{i=1}^{L} \left([s_i'(X)]^2 - [R_i - s_i(X)] s_i''(X) \right) \qquad (19)$$

$$\frac{\partial^2 \ln p(Y|X)}{\partial X^2} = -\frac{\sigma_Y^2}{\sigma_N^2\sigma_X^2} \qquad (20)$$

$$\frac{\partial^2 \ln p(X)}{\partial X^2} = -\frac{1}{\sigma_X^2}. \qquad (21)$$

Only (19) is stochastic, the expected value with respect to $\boldsymbol{R}$, Y, and X can be evaluated as follows

$$\begin{aligned}
&\mathbf{E}_{\boldsymbol{R}YX}\left[\frac{\partial^2 \ln p(\boldsymbol{R}|X)}{\partial X^2}\right] \\
&= -\frac{1}{\sigma_V^2}\sum_{i=1}^{L}\left(\mathbf{E}_X[(s_i'(X))^2] - \mathbf{E}_{\boldsymbol{R}X}[(R_i - s_i(X))s_i''(X)]\right) \\
&= -\frac{1}{\sigma_V^2}\sum_{i=1}^{L}\left(\mathbf{E}_X[(s_i'(X))^2] - \mathbf{E}_{V_iX}[V_i s_i''(X)]\right) \\
&= -\frac{1}{\sigma_V^2}\sum_{i=1}^{L}\left(\mathbf{E}_X[(s_i'(X))^2] - \mathbf{E}_{V_i}[V_i]\mathbf{E}_X[s_i''(X)]\right) \\
&= -\frac{1}{\sigma_V^2}\sum_{i=1}^{L}\mathbf{E}_X[(s_i'(X))^2]. \qquad (22)
\end{aligned}$$

(15) together with (20)–(22) finally give the expression for the CRLB as stated in (7).

Appendix B Optimized Parameters

Tables 1 and 2 show some good choices of parameters for the different mappings. The values have been found by a numerical grid search.

Table 1: Parameters for Mapping 1

k	SNR = 5 dB	SNR = 15 dB	SNR = 25 dB
$\rho = 0.9$	2.1	1.8	1.6
$\rho = 0.99$	6.3	5.5	4.9

Table 2: Parameters for Mapping 2

(k, m, r)	SNR = 5 dB	SNR = 15 dB	SNR = 25 dB
$\rho = 0.9$	$(2.0, 2.5, 0.2)$	$(1.8, 4.5, 0.35)$	$(1.6, 10.5, 0.4)$
$\rho = 0.99$	$(6.3, 2.5, 0.2)$	$(5.6, 4.5, 0.35)$	$(4.8, 10.5, 0.4)$

References

[1] A. D. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. on Information Theory*, vol. 22, no. 1, pp. 1–10, January 1976.

[2] D. Slepian and J. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. on Information Theory*, vol. 19, no. 4, pp. 471–480, July 1973.

[3] H. L. V. Trees, *Detection, Estimation, and Modulation Theory. Part I.* Wiley, 1968.

[4] J. M. Wozencraft and I. M. Jacobs, *Principles of Communication Engineering.* Wiley, 1965.

[5] N. Wernersson, M. Skoglund, and T. Ramstad, "Polynomial based analog source–channel codes," *IEEE Trans. on Communications*, vol. 57, no. 9, pp. 2600–2606, September 2009.

[6] P. A. Floor, T. A. Ramstad, and N. Wernersson, "Power constrained channel optimized vector quantizers used for bandwidth expansion," in *International Symposium on Wireless Communication Systems*, October 2007.

[7] V. Vaishampayan and S. I. R. Costa, "Curves on a sphere, shift-map dynamics, and error control for continuous alphabet sources," *IEEE Trans. on Information Theory*, vol. 49, no. 7, pp. 1658–1672, July 2003.

Paper E

Iterative Source–Channel Coding Approach to Witsenhausen's Counterexample

Johannes Kron, Ather Gattami,
Tobias J. Oechtering, and Mikael Skoglund

Submitted to *Automatica*.

Iterative Source–Channel Coding Approach to Witsenhausen's Counterexample

Johannes Kron, Ather Gattami,
Tobias J. Oechtering, and Mikael Skoglund

Abstract

In 1968, Witsenhausen introduced his famous counterexample where he showed that even in the simple linear quadratic static team decision problem, complex nonlinear decisions could outperform any given linear decision. This problem has served as a benchmark problem for decades where researchers try to achieve the optimal solution. This paper introduces a systematic iterative source–channel coding approach to solve problems of the Witsenhausen Counterexample-character. The advantage of the presented approach is its simplicity. Also, no assumptions are made about the shape of the space of policies. The minimal cost obtained using the introduced method is **0.16692462***, which is the lowest known thus far.*

1 Introduction

The most fundamental problem in control theory, namely the static output feedback problem has been open since the birth of control theory. The question is whether there is an efficient algorithm that can decide existence and find stabilizing controllers, linear or nonlinear, based on imperfect measurements and given memory. The static output feedback problem is just an instance of the problem of control with information structures imposed on the controllers, which has been very challenging for decision theory researchers. In 1968, Witsenhausen [21] introduced his famous counterexample:

$$\inf_{\gamma_1(\cdot),\gamma_2(\cdot)} \mathbf{E}\left[k^2\gamma_1^2(X_0) + X_2^2\right] \tag{1}$$

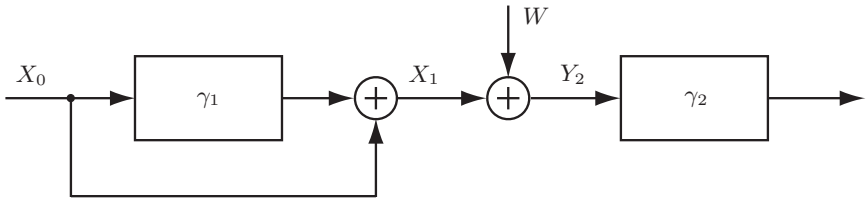


Figure 1: Schematic view of the system.

where

$$X_1 = \gamma_1(X_0) + X_0, \tag{2}$$
$$X_2 = X_1 - \gamma_2(Y_2), \tag{3}$$
$$Y_1 = X_0, \tag{4}$$
$$Y_2 = X_1 + W, \tag{5}$$

$X_0 \sim N(0, \sigma^2)$, and $W \sim N(0, 1)$. Here we have two decision makers, one corresponding to γ_1 and the other to γ_2. The problem is a two-stage linear quadratic Gaussian control problem, where the cost at the first time-step is $\mathbf{E}[k^2\gamma_1^2(X_0)]$ and $\mathbf{E}X_2^2$ at the second one. At the first time-step, the controller has full state measurement, $Y_1 = X_0$. At the second time-step, it has imperfect state measurement, $Y_2 = X_1 + W$. What is different to the classical output feedback problem, is that the controller at the second stage does not have information from the past since it has no information about the output Y_1. Thus, the controller is restricted to be a static output feedback controller. Witsenhausen showed that even in the simple linear quadratic Gaussian control problem above, complex nonlinear decisions could outperform any given linear decision. This problem has served as a benchmark problem for decades where researchers try to achieve the optimal solution. It has been pointed out that the problem is complicated due to a so called "signaling-incentive", where decisions are not only chosen to minimize a given cost, but also to encode information in the decisions in order to signal information to other decision makers in the team. In the example above, decision maker 2 measures $Y_2 = X_0 + \gamma_1(X_0) + W$, so its measurement is affected by decision maker 1 through γ_1. Hence, decision maker 1 not only tries to optimize the quadratic cost in (1), but also *signal* information about X_0 to decision maker 2 through its decision, $\gamma_1(X_0)$.

The problem also has a nice communication theoretic analogue. First, write Witsenhausen's counterexample as minimizing the cost

$$\mathbf{E}\left[k^2\gamma_1^2(X_0) + (X_0 + \gamma_1(X_0) - \gamma_2(X_0 + \gamma_1(X_0) + W))^2\right]$$

with respect to $\gamma_1(\cdot)$ and $\gamma_2(\cdot)$. Now consider the slightly modified problem

$$\begin{aligned} &\text{minimize } \mathbf{E}\,(X_0 - \gamma_2(\gamma_1(X_0) + W))^2 \\ &\text{subject to } \mathbf{E}\,\gamma_1^2 \leq p \end{aligned}$$

The modification made is that we removed X_0 term from the measurement of γ_2, and removed γ_1 from the objective function, and instead added a constraint $\mathbf{E}\,\gamma_1^2 \leq p$ to make sure that it has a limited variance (of course we could set an arbitrary power limitation on the variance). The modified problem is exactly the Gaussian channel coding/decoding problem!

Previous work has been pursued on understanding the Witsenhausen Counterexample. Suboptimal solutions where found in [13] studied variations of the problem when the signaling incentive was eliminated. In [12,14], connections to information theory where studied. An extensive study of the information theoretic connection was made in [3], where it was shown that coupling between decision makers in the cost function introduced the nonlinear behavior of the optimal strategies. An ordinal optimization approach was introduced in [5] and a hierarchical search approach was introduced in [16], where both rely on a given structure of the decisions. The first method that showed that optimal strategies may have "slopes" to the quantizations was given in [2]. Solutions with bounds are studied in [11]. A potential games approach in the paper by [17] found the best known value to the date of its publication, namely 0.1670790.

In this paper, we will introduce a generic method of iterative optimization based on ideas from source–channel coding [8, 9, 15, 20], that could be used to solve problems of the Witsenhausen Counterexample character. The numerical solution we obtain for the benchmark problem is of high accuracy and renders the lowest value known thus far, **0.16692462**.

2 Notation

$p(\cdot)$ and $p(\cdot|\cdot)$ denote probability density functions (pdfs) and conditional pdfs, respectively.

3 Iterative Optimization

We will now present an iterative design algorithm, based on person-by-person optimality, for solving the minimization in equation (1). The method we propose is related to the Lloyd–Max algorithm [10,18,19] that is successfully used when designing quantizers. A quantizer can be described by its partition cells and their corresponding reproduction value. The partition cells define to which codeword analog values are encoded and the reproduction values define how the analog value is reproduced from the codeword.

In general, there is no explicit, closed-form solution to the problem of finding the optimal quantizer [10]. The idea of the Lloyd–Max algorithm is to assume that either the partition cells or the reproduction values are fixed; with one part fixed, it is straightforward to derive an optimal expression for its counterpart. By optimizing one part at a time in an iterative fashion, the quantizer will converge to a local optimum. The Lloyd–Max algorithm has been generalized and used in various joint source–channel coding applications, for example, [6–8, 15, 20, 22]. The generalization of the Lloyd–Max algorithm that will be used in this paper involves four key elements:

1. Formulation of necessary conditions on γ_1 and γ_2 such that they are individually optimal given that γ_2 and γ_1, respectively, are fixed.

2. Discretization of the "channel" space between γ_1 and γ_2 such that X_1 and the input to γ_2 are restricted to belong to a finite set $\mathcal{S}_L$.

3. Iterative optimization of γ_1 and γ_2 to make sure that they, one at a time, fulfill their corresponding necessary conditions.

4. Use of a technique called noisy "channel" relaxation that makes the solution less sensitive to the initialization.

3.1 Necessary Conditions on γ_1

Let us first define the function $\tilde{\gamma}_1$ as

$$\tilde{\gamma}_1(x_0) \triangleq \gamma_1(x_0) + x_0 = x_1. \tag{6}$$

Without loss of generality, we will optimize with respect to $\tilde{\gamma}_1$. The cost we want to minimize is given by

$$J \triangleq \mathbf{E}[k^2\gamma_1^2(X_0) + (X_1 - \gamma_2(Y_2))^2]. \tag{7}$$

Using Bayes' rule, the expected cost function can now be expressed as

$$\begin{aligned} J &= \int p(x_0, y_2|\tilde{\gamma}_1)\ F(x_0, \tilde{\gamma}_1(x_0), \gamma_2(y_2))\ \mathrm{d}x_0\mathrm{d}y_2 \\ &= \int p(x_0)p(y_2|x_0, \tilde{\gamma}_1)\ F(x_0, \tilde{\gamma}_1(x_0), \gamma_2(y_2))\ \mathrm{d}x_0\mathrm{d}y_2 \\ &= \int p(x_0)p(y_2|\tilde{\gamma}_1(x_0))\ F(x_0, \tilde{\gamma}_1(x_0), \gamma_2(y_2))\ \mathrm{d}x_0\mathrm{d}y_2, \end{aligned} \tag{8}$$

where

$$F(x_0, x_1, \gamma_2(y_2)) = \Big(k^2(x_1 - x_0)^2 + (x_1 - \gamma_2(y_2))^2\Big). \tag{9}$$

Since the integrand in (8) is positive for all values of x_0, it is clear that the optimization of $\tilde{\gamma}_1$ (assuming γ_2 is fixed) can be done individually for each x_0. A necessary condition for $\tilde{\gamma}_1$ to be optimal is given by

$$\tilde{\gamma}_1(x_0) = \arg\min_{x_1 \in \mathbb{R}} \Big(\int p(y_2|x_1)\ F(x_0, x_1, \gamma_2(y_2))\ \mathrm{d}y_2 \Big) \tag{10}$$

for all $x_0 \in \mathbb{R}$.

3.2 Necessary Conditions on γ_2

If we now assume that γ_1 is fixed, we see that the the first term in (7) is a constant. The minimization of J with respect to γ_2 is therefore equivalent to

$$\min_{\gamma_2(\cdot)} \mathbf{E}[(X_1 - \gamma_2(Y_2))^2], \tag{11}$$

which is the mean-squared error (MSE). It is well known that the MSE is minimized by the conditional expected value; hence,

$$\gamma_2(y_2) = \mathbf{E}[X_1|y_2] \tag{12}$$

for all $y_2 \in \mathbb{R}$, is a necessary condition for $\gamma_2(y_2)$ to be optimal.

3.3 Discretization

Although (10) and (12) would be possible to numerically evaluate for a particular x_0 and y_2, respectively, they are impractical since the full representation of the functions is infinite-dimensional. To get around this problem we introduce the set

$$\mathcal{S}_L = \Big\{ -\Delta\frac{L-1}{2}, -\Delta\frac{L-3}{2}, \ldots, \Delta\frac{L-3}{2}, \Delta\frac{L-1}{2} \Big\}, \tag{13}$$

where $L \in \mathbb{N}$ and $\Delta \in \mathbb{R}_+$ are two parameters that determine the number of points and the spacing between the points, respectively. Next, we impose the constraint $x_1 \in \mathcal{S}_L$, that is, the output of $\tilde{\gamma}_1$ can only take one out of a finite number of values. In a similar way, the input to γ_2 is discretized such that,

$$\gamma_2(y_2) = \tilde{\gamma}_2(\tilde{y}_2), \quad \tilde{y}_2 = Q_{\mathcal{S}_L}(y_2) \in \mathcal{S}_L, \tag{14}$$

where $Q_{\mathcal{S}_L}(y_2)$ maps y_2 to the closest point in the set $\mathcal{S}_L$. γ_2 can now be stored in the form of a lookup table where each point in $\mathcal{S}_L$ is associated with an output value. The approximation of the real space with $\mathcal{S}_L$ can be

made more and more accurate by decreasing Δ and increasing L.[1] Finally, since X_0 is still infinite-dimensional, we use Monte-Carlo samples of X_0 to represent the input to $\tilde{\gamma}_1$. $\tilde{\gamma}_1$ is now specified by evaluating

$$\tilde{\gamma}_1(x_0) = \arg \min_{x_1 \in \mathcal{S}_L} \sum_{\tilde{y}_2 \in \mathcal{S}_L} p(\tilde{y}_2|x_1)\ F(x_0, x_1, \tilde{\gamma}_2(\tilde{y}_2)) \tag{15}$$

for each of the Monte-Carlo samples that represent X_0. In a similar way, $\tilde{\gamma}_2$ can be expressed as

$$\tilde{\gamma}_2(\tilde{y}_2) = \mathbf{E}[X_1|\tilde{y}_2], \tag{16}$$

for all $\tilde{y}_2 \in \mathcal{S}_L$, where the expectation with respect to X_0 is evaluated by using the Monte-Carlo samples.

3.4 Design Algorithm

Given the above expressions for $\tilde{\gamma}_1$ and $\tilde{\gamma}_2$ it will be possible to optimize the system iteratively. We do this by keeping one part of the system fixed while we optimize the other part. One common problem with iterative techniques like the one suggested here is that the final solution will depend on the initialization of the algorithm. If the initialization is bad we are likely to end up in a poor local minimum.

In joint source–channel coding, one method that has proven to be helpful in counteracting this is noisy channel relaxation (NCR) [8, 9, 15, 20]. The idea of NCR is to change some parameter and first design a system for a completely different scenario with a simpler solution. The solution that is obtained is then used as initialization when designing for a scenario that is a bit closer to the true scenario. In joint source–channel coding, this is done by first designing a system for a channel with a lower signal-to-noise ratio (SNR) than the target SNR, which explains the name of the method. In the Witsenhausen setup, we have found that the ideas from NCR can be used as follows. Design a system for a high value of k first and then gradually decrease k until the desired value of k is reached. The reason to start with a high value of k is that the design algorithm will find a solution where $\tilde{\gamma}_1(x_0) \approx x_0$ in this case (i.e., $\gamma_1(x_0) \approx 0$) independently of γ_2. The design procedure including the NCR part is given in Algorithm 1.

Each update on line 7 and 8 in Algorithm 1 will decrease the cost. Since the cost is lower bounded, it is clear that the algorithm will converge. It may happen that the algorithm converges to a local optimum, however, as will be seen in the following section the *local* optima we obtain are still better than any previously reported results.

[1] While decreasing Δ, one has to increase L to make sure that $\max(x \in \mathcal{S}_L) = \Delta(L-1)/2$ does not decrease.

Algorithm 1 Design Algorithm

Require: Initial mapping of $\tilde{\gamma}_2$, the value k for which the system should be optimized and the threshold δ that determines when to stop the iterations.

Ensure: Locally optimized $\tilde{\gamma}_1$ and $\tilde{\gamma}_2$.

1: Let $k' > k$.
2: **while** $k' > k$ **do**
3: Decrease k' according to some scheme (e.g., linearly).
4: Set the iteration index $i = 0$ and $J^{(0)} = \infty$.
5: **repeat**
6: Set $i = i + 1$
7: Find the optimal $\tilde{\gamma}_1$ by using (15).
8: Find the optimal $\tilde{\gamma}_2$ by using (16).
9: Evaluate the cost function $J^{(i)}$ according to (7).
10: **until** $(J^{(i-1)} - J^{(i)})/J^{(i-1)} < \delta$
11: **end while**

4 Results

4.1 Implementation Aspects

For the evaluation of the design algorithm we have initially used $L = 201$ levels and chosen $\Delta(L) = 10\sigma/(L-1)$. We have used 400000 Monte-Carlo samples in the final optimizations to represent X_0. Since it is known that the optimal γ_1 is symmetric about origin [21], we have restricted $\tilde{\gamma}_1$ to have this symmetry by generating only positive Monte-Carlo samples and thereafter reflecting the resulting $\tilde{\gamma}_1$-function for negative values of x_0.

To be able to compare our results to previously reported results, we have set $\sigma = 5$ and $k = 0.2$. However, since we are using the NCR idea, we have initially used the value $k' = 3$ and decreased it according to the series $\{3, 2, 1.5, 1, 0.6, 0.4, 0.3, 0.2\}$. Before running the design algorithm, we require $\tilde{\gamma}_2$ to be initialized. However, due to the NCR this has little impact on the final solution and we have used the initialization $\tilde{\gamma}_2 \equiv 0$.

Once we have obtained the solution for $k' = 0.2$, we have increased the precision by expanding the number of points in the discrete set from L to L' and updated $\tilde{\gamma}_2$ according to

$$\tilde{\gamma}_2^{(L')}(\tilde{y}_2) = \tilde{\gamma}_2^{(L)}(Q_{\mathcal{S}_L}(\tilde{y}_2)) \tag{17}$$

for all $\tilde{y}_2 \in \mathcal{S}_{L'}$. Thereafter the inner part of the design algorithm, that is, lines 4–10, have been run again to obtain a system optimized for the increased number of points L'. By repeating this refinement, the precision increases and the cost decreases as will be shown later. This method of

Table 1: Final cost for different solutions ($L = 12801$)

Steps	Stage-1 Cost	Stage-2 Cost	Total Cost
Witsenhausen [21][a]	0.40423088	0.00002232	0.40425320
Bansal and Bansar [3][a]	0.36338023	0.00163460	0.36501483
Deng and Ho [5][a]	0.13948840	0.05307290	0.19256130
Baglietto et al. [2]			0.1701
Lee et al. [16]	0.13188408	0.03542912	0.16731321
Li et al. [17]			0.1670790
This paper, 3-step	0.13493778	0.03201113	0.16694891
This paper, 3.5-step	0.13462186	0.03230369	0.16692555
This paper, 4-step	0.13484828	0.03207634	**0.16692462**

[a] Costs obtained from [16].

refining the precision is similar to the one-way multigrid algorithm that is analyzed in [4]. The evaluations of (15) and (16) have been done using an exhaustive search, therefore, the run time is exponential in the number of levels L. An example of the number of iterations that are required and the total computation time can be found in Appendix B.

4.2 Numerical Results

During the first steps of the NCR k' is high. This means that the output of $\tilde{\gamma}_1$ should follow the input closely to avoid large costs in the first stage. If continuous outputs were allowed, the output would be identical to the input. However, since we are working with a discretized system, only outputs from the set $\mathcal{S}_L$ are possible. As k' reaches 0.4–0.6 the step behavior of the output appears. Depending on the realization of the Monte-Carlo samples we get either a 3.5-step mapping as shown in Figure 2 or a 4-step mapping as shown in Figure 3 (occasionally, a 3-step solution has occurred). The total costs for these solutions are stated in Table 1. For ease of comparison, we have also included the costs of previously reported results. As can be seen, all our mappings have similar performance and all of them give lower costs than the previously reported lowest cost — 0.1670790 [17].

In Table 2 we show how the cost decreases as the number of points L is increased. The method we use to calculate the total cost as well as some notes on the accuracy can be found in Appendix A. The lowest cost we have achieved with our algorithm is 0.16692462. The mapping that achieves this cost is the 4-step mapping shown in Figure 3 with $L = 12801$ points. Although the mapping contains four clear output levels it should be emphasized that each level is slightly sloped; this can be seen in Figure 4, where the first step has been zoomed in. It is reasonable to assume that

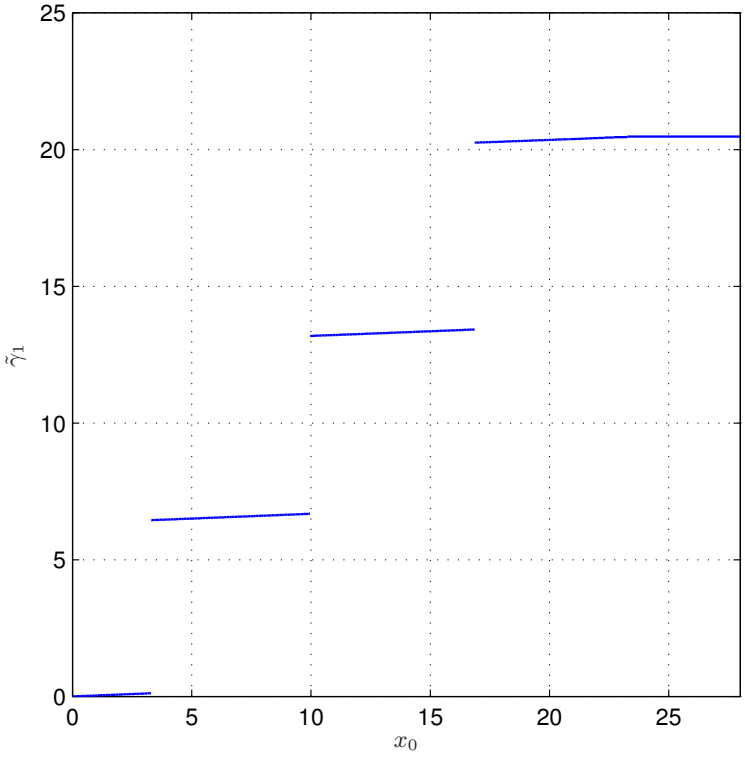


Figure 2: 3.5-step solution ($L = 12801$)

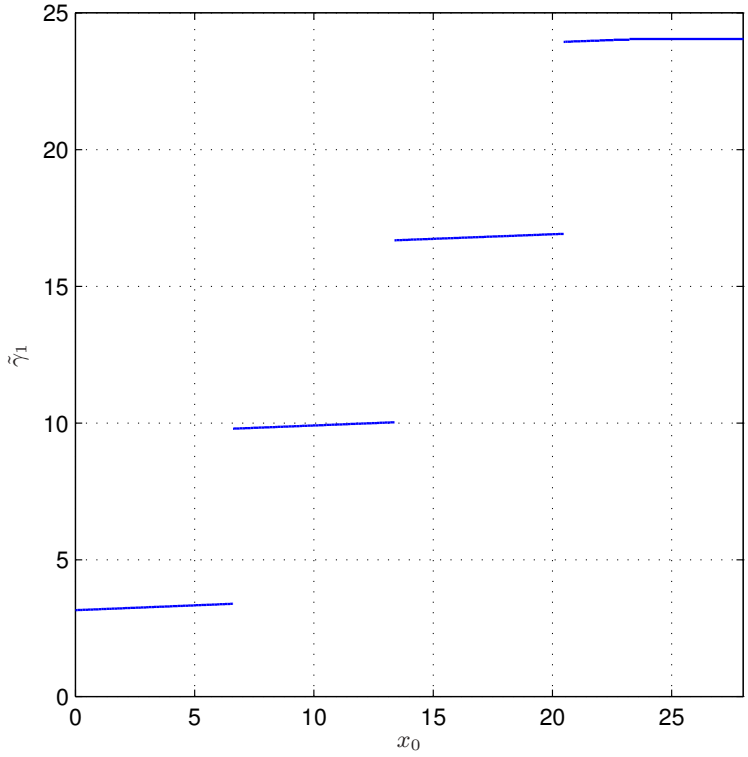


Figure 3: 4-step solution ($L = 12801$)

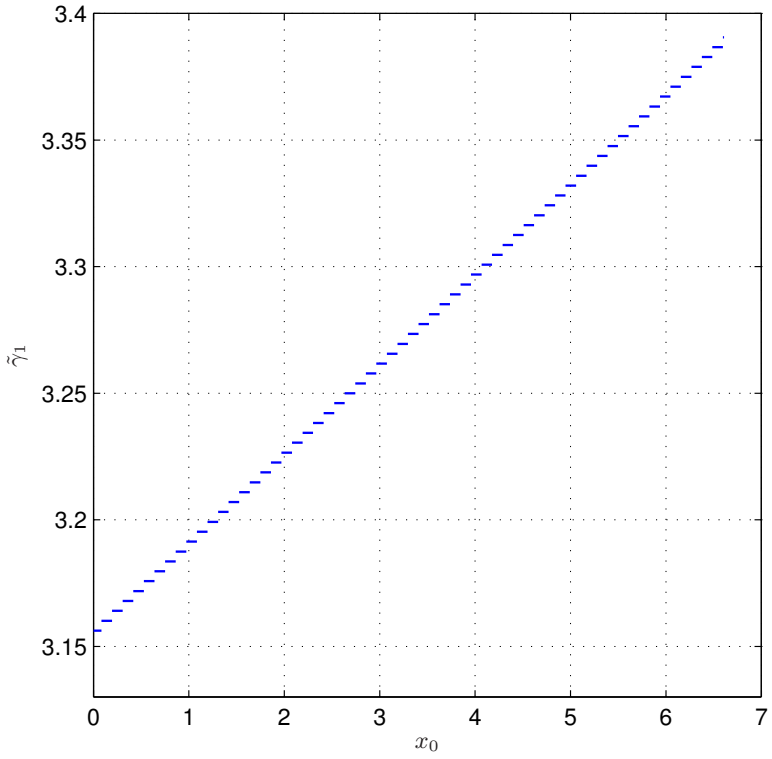


Figure 4: Detailed view of the first step in the 4-step solution.

Table 2: Costs for different precisions for the 4-step solution

L	M	Stage-1 Cost	Stage-2 Cost	Total Cost
201	16	0.12104248	0.05764052	0.17868301
401	22	0.13015033	0.03883388	0.16898421
801	30	0.13530785	0.03200857	0.16731642
1601	56	0.13496610	0.03206243	0.16702853
3201	110	0.13486814	0.03208140	0.16694954
6401	210	0.13485895	0.03207070	0.16692966
12801	396	0.13484828	0.03207634	0.16692462

as the precision (i.e., L) increases further, each step of the mapping will converge to a straight line that is slightly sloped.

5 Comparison to Previous Results

In this section, we will compare the presented method with the previous methods. Besides the fact that the new method improves the optimized cost, there are further advantages compared to previous work:

- No structure is assumed for the decision functions. In [5] and [16], monotonicity of the decisions was assumed. The space of decisions is assumed to be a normed linear space in [2].
- A significant analytic/modeling work was performed before posing the optimization problem to be solved in [5], [16], and [2]. The first two require manual adjustments for the proper choice of interval values and signal levels, and the third requires some prior analysis to determine a constant "c". In [17], modeling work is needed in converting the problem into a potential game.

The method presented in this paper is fully automated and its advantages rely on the following:

- Discretization of the decision space.
- Iterative optimization between the different decision functions.
- A noisy channel relaxation technique that makes the solution insensitive to the initialization.

6 Conclusions

In this paper, we introduced a generic method of iterative optimization based on ideas from Source–Channel coding, that could be used to solve

problems of the Witsenhausen Counterexample character. The numerical solution we obtain for the benchmark problem is of high accuracy and renders the lowest value known thus far, **0.16692462**.

Also, the design algorithm does not make any assumption on the structure of the policies — the solutions are allowed to have arbitrary shapes (within the restrictions imposed by the discretization). The results can therefore be seen as a confirmation that the step-shaped behavior is beneficial.

Appendix A Calculation of the Total Cost

In the design algorithm, $\tilde{\gamma}_1$ is specified implicitly by, for each Monte-Carlo sample, storing the output symbol to which it is mapped. This representation is used when evaluating the cost during the iterations in the design algorithm. However, to evaluate the final total cost we need higher numerical accuracy. Therefore, the first step in calculating the total cost is to use the sample-based representation to find an explicit specification of $\tilde{\gamma}_1$ given by

$$\tilde{\gamma}_1(x_0) = \alpha_i \in \mathcal{S}_L \quad \text{if } A_i \leq x_0 < A_{i+1}, \tag{18}$$

for $i \in \{0, \ldots, M-1\}$, with $A_0 = -\infty$ and $A_M = \infty$. That is, we transform the sample-based representation of $\tilde{\gamma}_1$, which is explicitly defined only for the Monte-Carlo samples, to a function which is defined for all real numbers. This representation makes it possible to numerically evaluate the integrals that are needed to find the total cost

$$\begin{aligned} J &= \mathbf{E}[k^2\gamma_1^2(X_0) + (X_1 - \gamma_2(Y_2))^2] \\ &= \underbrace{\mathbf{E}[k^2(\tilde{\gamma}_1(X_0) - X_0)^2]}_{=J_1} + \underbrace{\mathbf{E}[(\tilde{\gamma}_1(X_0) - \tilde{\gamma}_2(\tilde{Y}_2))^2]}_{=J_2}, \end{aligned} \tag{19}$$

where

$$\begin{aligned} J_1 &= \int_{x_0} p(x_0)k^2(\tilde{\gamma}_1(x_0)-x_0)^2\mathrm{d}x_0 \\ &= k^2\sum_{i=0}^{M-1}\int_{A_i}^{A_{i+1}} p(x_0)(\alpha_i - x_0)^2\mathrm{d}x_0, \end{aligned} \tag{20}$$

$$\begin{aligned} J_2 &= \int_{x_0}\sum_{\tilde{y}_2\in\mathcal{S}_L} p(x_0,\tilde{y}_2)(\tilde{\gamma}_1(x_0)-\tilde{\gamma}_2(\tilde{y}_2))^2\mathrm{d}x_0 \\ &= \int_{x_0}\sum_{\tilde{y}_2\in\mathcal{S}_L} p(x_0)P(\tilde{y}_2|x_0)(\tilde{\gamma}_1(x_0)-\tilde{\gamma}_2(\tilde{y}_2))^2\mathrm{d}x_0 \\ &= \sum_{i=0}^{M-1}\int_{A_i}^{A_{i+1}} p(x_0)\sum_{\tilde{y}_2\in\mathcal{S}_L} P(\tilde{y}_2|\alpha_i)(\alpha_i-\tilde{\gamma}_2(\tilde{y}_2))^2\mathrm{d}x_0 \\ &= \sum_{i=0}^{M-1}\Big\{\sum_{\tilde{y}_2\in\mathcal{S}_L} P(\tilde{y}_2|\alpha_i)(\alpha_i-\tilde{\gamma}_2(\tilde{y}_2))^2\Big\}\int_{A_i}^{A_{i+1}} p(x_0)\mathrm{d}x_0, \end{aligned} \tag{21}$$

and

$$P(\tilde{y}_2|\alpha_i) = \begin{cases} \int_{-\infty}^{\tilde{y}_2+\Delta/2} p(w=y_2-\alpha_i)\mathrm{d}y_2 & \text{if } \tilde{y}_2 = -\Delta\frac{L-1}{2} \\ \int_{\tilde{y}_2-\Delta/2}^{\infty} p(w=y_2-\alpha_i)\mathrm{d}y_2 & \text{if } \tilde{y}_2 = \Delta\frac{L-1}{2} \\ \int_{\tilde{y}_2-\Delta/2}^{\tilde{y}_2+\Delta/2} p(w=y_2-\alpha_i)\mathrm{d}y_2 & \text{otherwise} \end{cases} \tag{22}$$

All integrals have been calculated numerically using the Matlab function *quadl* with the tolerance specified to be $t = 10^{-18}$, which means that the absolute error of the result from quadl is not greater than t. All integrands are continuous and have a smooth behavior that should cause no problem for quadl. To upper bound the total cost, we have upper bounded each integral by adding t to each individual result from quadl and reevaluated the total cost. In this way we have estimated the absolute error to be in the order of (or less than) 10^{-11}. Matlab code for our calculations of the total cost, including our decision functions can be found in [1].

Appendix B Computation time

Table 3 and 4 show the computation times for the design algorithm and the refinement process, respectively. The algorithms were implemented in C++ on a computer with an Intel Core2 Quad CPU running at 2.66 GHz.

Table 3: Computation time — Algorithm 1 ($L = 201$)

k'	# Iterations	Time
3	4	32 s
2	2	13 s
1.5	2	14 s
1	2	13 s
0.6	2	13 s
0.4	100	11 min
0.3	36	4 min
0.2	30	3 min
Total	178	20 min

Table 4: Computation time — Refinement

L	# Iterations	Time
401	44	15 min
801	83	98 min
1601	12	52 min
3201	3	49 min
6401	2	126 min
12801	2	501 min
Total		14 h

References

[1] `http://www.ee.kth.se/~johk/witsenhausen/`.

[2] M. Baglietto, T. Parisini, and R. Zoppoli. Numerical solutions to the Witsenhausen counterexample by approximating networks. *IEEE Trans. Automatic Control*, 46(9):1471–1477, September 2001.

[3] R. Bansal and T. Basar. Stochastic teams with nonclassical information revisited: When is an affine law optimal? *IEEE Trans. Automatic Control*, 32(6):554–559, June 1987.

[4] C.S. Chow and J. N. Tsitsiklis. An optimal one-way multigrid algorithm for discrete-time stochastic control. *Automatic Control, IEEE Transactions on*, 36(8):898–914, 1991.

[5] M. Deng and Y.C. Ho. An ordinal optimization approach to optimal control problems. *Automatica*, 35:331–338, 1999.

[6] N. Farvardin and V. Vaishampayan. Optimal quantizer design for noisy channels: An approach to combined source–channel coding. *IEEE Trans. on Information Theory*, 33(6):827–838, November 1987.

[7] N. Farvardin and V. Vaishampayan. On the performance and complexity

of channel-optimized vectorquantizers. *IEEE Trans. on Information Theory*, 37(1):155–160, January 1991.

[8] A. Fuldseth and T. A. Ramstad. Bandwidth compression for continuous amplitude channels based on vector approximation to a continuous subset of the source signal space. In *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 3093–3096, Munich, Germany, April 1997.

[9] S. Gadkari and K. Rose. Noisy channel relaxation for VQ design. In *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 2048–2051, May 1996.

[10] A. Gersho and R. M. Gray. *Vector Quantization and Signal Compression*. Kluwer academic publishers, Dordrecht, The Netherlands, 1992.

[11] P. Grover, S. Y. Park, and A. Sahai. The finite-dimensional Witsenhausen counterexample. In *ConCom*, Seoul, Korea, March 2009.

[12] Y.C. Ho and T. S. Chang. Another look at the nonclassical information structure problem. *IEEE Trans. on Automatic Control*, 25(3), 1980.

[13] Y.C. Ho and K.C. Chu. Team decision theory and information structures in optimal control problems—part I. *IEEE Trans. on Automatic Control*, 17(1), 1972.

[14] Y.C. Ho, M. P. Kastner, and E. Wong. Teams, signaling, and information theory. *IEEE Trans. on Automatic Control*, 23(2), 1978.

[15] J. Karlsson and M. Skoglund. Optimized low-delay source–channel–relay mappings. *IEEE Trans. on Communications*, 58(5):1397–1404, May 2010.

[16] J. T. Lee, E. Lau, and Y.C. Ho. The Witsenhausen counterexample: A hierachical search approach for nonconvex optimization problems. *IEEE Trans. Automatic Control*, 46(3):382–297, March 2001.

[17] N. Li, R. Marden, and J. S. Shamma. Learning approaches to the Witsenhausen counterexample from a view of potential games. In *IEEE Conference on Decision and Control*, pages 157–162, December 2009.

[18] S. P. Lloyd. Least Squares Quantization in PCM. *IEEE Trans. on Information Theory*, 28(2):129–137, March 1982.

[19] J. Max. Quantizing for minimum distortion. *IRE Trans. on Information Theory*, 6:7–12, March 1960.

[20] N. Wernersson, J. Karlsson, and M. Skoglund. Distributed quantization over noisy channels. *IEEE Trans. on Communications*, 57(6):1693–1700, June 2009.

[21] H. S. Witsenhausen. A counterexample in stochastic optimum control. *SIAM Journal on Control*, 6(1):138–147, 1968.

[22] K. A. Zeger and A. Gersho. Vector quantizer design for memoryless noisy channels. In *IEEE Internationell Conference on Communications*, pages 1593–1597, Philadelphia, USA, June 1988.

Paper F

Optimized Low-Delay Source–Channel–Relay Mappings

Johannes Karlsson and Mikael Skoglund

Published in *IEEE Transactions on Communications*, May 2010.

Optimized Low-Delay Source–Channel–Relay Mappings

Johannes Karlsson and Mikael Skoglund

Abstract

The three-node relay channel with a Gaussian source is studied for transmission subject to a low-delay constraint. A design algorithm for joint source–channel mappings is proposed and numerically evaluated. The designed system is compared with reference systems, based on modular source and channel coding, and the distortion-rate function for the Gaussian source using known achievable rates for the relay channel. There is a significant gain, in terms of decreased power, in using the (locally) optimized systems compared with the reference systems. The structure of the resulting source mapping and the relay mapping is visualized and discussed in order to gain understanding of fundamental properties of optimized systems. Interestingly, the design algorithm generally produces relay mappings with a structure that resembles Wyner–Ziv compression.

1 Introduction

The relay channel has been studied extensively since its introduction [1]. With the increasing popularity and relevance of ad-hoc wireless sensor networks, cooperative transmission is more relevant than ever. In this paper, we focus on relaying in the context of source transmission over a sensor network. A sensor node encodes measurements and communicates these to a sink node, with another node acting as a relay in the transmission. We focus on low-delay memoryless source–channel and relay mappings, subject to power constraints at the source and relay nodes. Hence, the proposed technique is a suitable candidate in applications with strict delay and energy constraints, such as in wireless sensor networking for closed-loop control over wireless channels [2, 3].

Existing work on source and channel coding over the relay channel includes [4, 5]. However, whereas [4] looks at asymptotic high-SNR properties

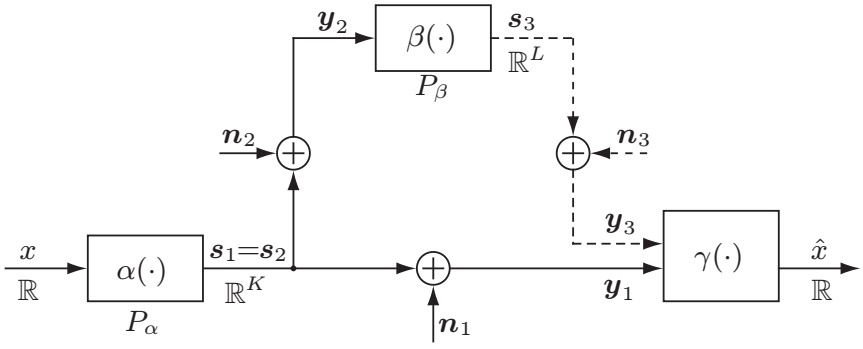


Figure 1: Structure of the system.

the present work is design oriented. Also, although [5] includes some practical results it relies on powerful channel codes. Because of this, the decoding is not instantaneous but a significant delay is needed for the message to be decoded. Another recent study is the one presented in [6]. This work also focuses on characterizing the achievable high-SNR performance, however, in the presence of partial channel-state feedback.

The source–channel separation theorem [7] states that source and channel coding can be treated separately. However, in the case of *low-delay* constraints this is no longer true. We therefore propose a joint source–channel coding solution where at the source node, the source and channel codes are merged into one single operation — a mapping from the source space to the channel space. In a similar way, the operation at the relay is a mapping from its input channel space to its output channel space. We investigate how to optimize[1] both the source–channel mapping at the source as well as the channel–channel mapping at the relay. To our knowledge, there are no similar existing results in this direction. Our approach is however related to the ones being used for bandwidth compression–expansion in [8–10] and distributed source–channel coding in [11].

2 Problem Formulation

We will study the three-node system depicted in Figure 1. Our goal is to transmit information about the Gaussian random variable X with variance $\sigma_X^2 = 1$ from the source node to the destination node so that it can be reconstructed with the smallest possible distortion. Besides the direct link

[1] We use the term *optimized*, in contrast to *optimal*, to refer to a system which is locally optimal but not necessarily globally optimal.

we also have a path from the source to the destination via the relay node. The rules for the communication are the following. For each source sample X we have T channel uses at hand. The source and the relay do not transmit at the same time but must share these channel uses. We therefore use K channel uses for the transmission from the source and the remaining $L = T - K$ channel uses for the transmission from the relay. The scenario is in other words that of a half-duplex orthogonal relay channel. All transmissions are disturbed by additive white Gaussian noise; the received symbols on each channel can therefore be expressed as

$$\boldsymbol{y}_i = \boldsymbol{s}_i + \boldsymbol{n}_i \quad i = 1, 2, 3, \tag{1}$$

where $\boldsymbol{s}_i$ is the transmitted symbol and $\boldsymbol{n}_i$ is independent white Gaussian noise with $\mathbf{E}[\boldsymbol{n}_i \boldsymbol{n}_i^T] = \sigma_i^2 I$, $i = 1, 2, 3$. The transmitted symbols are given by the functions α and β according to

$$\boldsymbol{s}_1 = \boldsymbol{s}_2 = \alpha(x) \in \mathbb{R}^K, \tag{2}$$

$$\boldsymbol{s}_3 = \beta(\boldsymbol{y}_2) \in \mathbb{R}^L. \tag{3}$$

The equality $\boldsymbol{s}_1 = \boldsymbol{s}_2$ is due to the broadcast nature of a wireless channel. The source and the relay node operate under average transmit power constraints given by

$$\frac{1}{K}\mathbf{E}[\|\alpha(X)\|^2] \leq P_\alpha, \tag{4}$$

$$\frac{1}{L}\mathbf{E}[\|\beta(\boldsymbol{Y}_2)\|^2] \leq P_\beta. \tag{5}$$

For notational convenience we define the channel power gain of each channel as $a_i = 1/\sigma_i^2$. Assuming (4) is fulfilled with equality, the total signal-to-noise ratio (SNR) of the transmission from the source node to the destination node is hence given by $P_\alpha a_1$. We further assume that the total SNR of all channels is known by all involved parts. The destination node receives two symbols — $\boldsymbol{y}_1$ from the direct link and $\boldsymbol{y}_3$ from the relay. Based on these the transmitted value is estimated as

$$\hat{x} = \gamma(\boldsymbol{y}_1, \boldsymbol{y}_3). \tag{6}$$

Given this system we want to find the optimal source mapping, relay mapping, and receiver — denoted α, β, and γ. To have a low-delay system we want the source and the relay nodes to work on a sample-by-sample basis restricting K and L to be integers. If $K > 1$, α will in general be a nonlinear mapping from the one-dimensional source space to the K-dimensional channel space. In a similar way β will be a nonlinear mapping from the K-dimensional input of the relay to its L-dimensional output. As distortion measure we use the mean squared error (MSE), $\mathbf{E}[(X - \hat{X})^2]$, "optimal" therefore refers to optimal in the minimum MSE (MMSE) sense.

3 Optimized Mappings

The expected distortion for a given system can be written as

$$D = \mathbf{E}[(X - \hat{X})^2] = \iiiint p(x)p(\boldsymbol{y}_1|\alpha(x))p(\boldsymbol{y}_2|\alpha(x))\times$$
$$p(\boldsymbol{y}_3|\beta(\boldsymbol{y}_2))(x - \gamma(\boldsymbol{y}_1, \boldsymbol{y}_3))^2 \mathrm{d}x\mathrm{d}\boldsymbol{y}_1\mathrm{d}\boldsymbol{y}_2\mathrm{d}\boldsymbol{y}_3, \tag{7}$$

where $p(\cdot)$ and $p(\cdot|\cdot)$ denote probability density functions (pdfs) and conditional pdfs, respectively. The factorization of $p(x, \boldsymbol{y}_1, \boldsymbol{y}_2, \boldsymbol{y}_3)$ in (7) follows from the fact that all channels are orthogonal with independent noises. What we would like is to find α, β, and γ such that D is minimized given the power constraints in (4) and (5). There are two problems with this direct approach. First, it is very hard to optimize all parts of the system simultaneously; second, the optimal mappings could be arbitrary nonlinear mappings with no closed form expressions. To make the problem feasible we take the following suboptimal approach. Instead of optimizing all parts of the system simultaneously we use the common strategy of optimizing one part at a time keeping the others fixed. The second problem is solved by discretizing each dimension of the channel space into M equally spaced points with spacing Δ according to

$$\mathcal{S} = \{-\Delta\frac{M-1}{2}, -\Delta\frac{M-3}{2}, \ldots, \Delta\frac{M-3}{2}, \Delta\frac{M-1}{2}\} \tag{8}$$

and restricting the outputs of the source and the relay node to satisfy $\boldsymbol{s}_1 \in \mathcal{S}^K$ and $\boldsymbol{s}_3 \in \mathcal{S}^L$, respectively. At the receiving side the same approximation is made using a hard decision decoding rule — for instance, $\boldsymbol{y}_1$ is decoded according to

$$\hat{\boldsymbol{y}}_1 = \arg \min_{\boldsymbol{y}_1' \in \mathcal{S}^K} \|\boldsymbol{y}_1 - \boldsymbol{y}_1'\|^2, \tag{9}$$

where " $\hat{}$ " will be used to indicate that the value has been discretized. This approximation is expected to be good as long as M is sufficiently large and Δ is small in relation to the standard deviation of the channel noise, σ_i. In the following analysis $P(\cdot|\cdot)$ will be used for conditional probabilities — for example, $P(\hat{\boldsymbol{y}}_3|\boldsymbol{s_1})$ denotes the probability that the relay receives $\hat{\boldsymbol{y}}_3$ given that $\boldsymbol{s}_1$ is transmitted from the source.

3.1 Optimal Source Mapping

The problem of finding the optimal source mapping α (assuming β and γ are fixed) is a constrained optimization problem, which can be turned into the following unconstrained problem using the Lagrange multiplier method [12, 13]

$$\min_{\alpha} \left(\mathbf{E}[(X - \hat{X})^2] + \lambda \mathbf{E}[\|\alpha(X)\|^2] \right), \tag{10}$$

where

$$\mathbf{E}[(X-\hat{X})^2] = \int p(x)\mathbf{E}[(x-\hat{X})^2|\alpha(x)]\mathrm{d}x, \tag{11}$$

$$\mathbf{E}[\|\alpha(X)\|^2] = \int p(x)\|\alpha(x)\|^2\mathrm{d}x. \tag{12}$$

Since $p(x)$ in (11)–(12) is nonnegative, it is clear that the operation of the source mapping, α, can be optimized for each x individually according to

$$\alpha(x) = \arg\min_{\boldsymbol{s}_1\in\mathcal{S}^K}\left(\mathbf{E}[(x-\hat{X})^2|\boldsymbol{s}_1] + \lambda\|\boldsymbol{s}_1\|^2\right) \tag{13}$$

where

$$\begin{aligned}\mathbf{E}[(x-\hat{X})^2|\boldsymbol{s}_1] = \quad & \textstyle\sum_{\hat{\boldsymbol{y}}_1,\hat{\boldsymbol{y}}_2,\hat{\boldsymbol{y}}_3} P(\hat{\boldsymbol{y}}_1|\boldsymbol{s}_1)P(\hat{\boldsymbol{y}}_2|\boldsymbol{s}_1)\times \\ & P(\hat{\boldsymbol{y}}_3|\beta(\hat{\boldsymbol{y}}_2))(x-\gamma(\hat{\boldsymbol{y}}_1,\hat{\boldsymbol{y}}_3))^2.\end{aligned} \tag{14}$$

The intuition behind the Lagrange term $\lambda\|\boldsymbol{s}_1\|^2$ is the following: $\|\boldsymbol{s}_1\|^2$ is a measure of the power that is needed to transmit the signal $\boldsymbol{s}_1$, the term $\lambda\|\boldsymbol{s}_1\|^2$ can therefore be used to control the transmit power of the source node by penalizing signals that would use too much power. When $\lambda \geq 0$ is set to the "correct" value, the source node will not map x to the signal that gives the lowest distortion but rather to the signal that gives the lowest distortion conditioned that the power constraint in (4) is fulfilled.

3.2 Optimal Relay Mapping

In a similar way, the minimization to find the optimal relay mapping β (assuming α and γ are fixed), can be turned into the following unconstrained minimization problem

$$\min_{\beta}\left(\mathbf{E}[(X-\hat{X})^2] + \eta\mathbf{E}[\|\beta(\hat{\boldsymbol{Y}}_2)\|^2]\right), \tag{15}$$

where

$$\mathbf{E}[(X-\hat{X})^2] = \sum_{\hat{\boldsymbol{y}}_2} P(\hat{\boldsymbol{y}}_2)\mathbf{E}[(X-\hat{X})^2|\hat{\boldsymbol{y}}_2,\beta(\hat{\boldsymbol{y}}_2)], \tag{16}$$

$$\mathbf{E}[\|\beta(\hat{\boldsymbol{Y}}_2)\|^2] = \sum_{\hat{\boldsymbol{y}}_2} P(\hat{\boldsymbol{y}}_2)\|\beta(\hat{\boldsymbol{y}}_2)\|^2. \tag{17}$$

Equations (11) and (16) are two different ways of expanding the MSE using Bayes' rule. Looking at (16) and (17), it is once again clear that the minimization can be done individually for each $\hat{\boldsymbol{y}}_2 \in \mathcal{S}^K$, which gives

$$\beta(\hat{\boldsymbol{y}}_2) = \arg\min_{\boldsymbol{s}_3\in\mathcal{S}^L}\left(\mathbf{E}[(X-\hat{X})^2|\hat{\boldsymbol{y}}_2,\boldsymbol{s}_3] + \eta\|\boldsymbol{s}_3\|^2\right) \tag{18}$$

where

$$\mathbf{E}[(X-\hat{X})^2|\hat{\boldsymbol{y}}_2,\boldsymbol{s}_3] = \textstyle\sum_{\hat{\boldsymbol{y}}_1,\hat{\boldsymbol{y}}_3} P(\hat{\boldsymbol{y}}_3|\boldsymbol{s}_3)\times \\ \int_x p(x|\hat{\boldsymbol{y}}_2)P(\hat{\boldsymbol{y}}_1|\alpha(x))(x-\gamma(\hat{\boldsymbol{y}}_1,\hat{\boldsymbol{y}}_3))^2\mathrm{d}x. \tag{19}$$

In (18), $\eta \geq 0$ is the Lagrange multiplier which — when chosen correctly — makes sure that the power constraint (5) is satisfied.

Sawtooth Mappings ($K = L = 1$): As we will see in Section 4, all of the optimized relay mappings have a similar shape in the one-dimensional case (i.e., $K = L = 1$). Based on this observation we propose to use a sawtooth mapping as shown in Figure 2. This mapping has previously been proposed for distributed source–channel coding [14] and also for the relay channel in the context of maximum achievable rates [15].

The sawtooth mapping can be parametrized by the two parameters b and c and is defined as

$$\beta(y_2) = \begin{cases} cy_2 & \text{if } y_2 \in [-b,b) \\ \beta(y_2-2bm) & \text{if } y_2-2bm \in [-b,b),\ m\in\mathbb{Z}, \end{cases} \tag{20}$$

where, for a given b, the parameter c must be chosen so that the power constraint in (5) is satisfied, that is, $\mathbf{E}[\beta^2(Y_2)] = P_\beta$. The optimal value of b will depend on the channel gains and is easiest found by performing a grid search.

3.3 Optimal Receiver

Since we use the MSE as a distortion measure, it is a well known fact from estimation theory that the optimal receiver (assuming α and β are fixed) is the expected value of X given the received symbols,

$$\hat{x} = \gamma(\hat{\boldsymbol{y}}_1,\hat{\boldsymbol{y}}_3) = \mathbf{E}[X|\hat{\boldsymbol{y}}_1,\hat{\boldsymbol{y}}_3] = \\ \frac{\displaystyle\int_x xp(x)P(\hat{\boldsymbol{y}}_1|\alpha(x))\sum_{\hat{\boldsymbol{y}}_2} P(\hat{\boldsymbol{y}}_2|\alpha(x))P(\hat{\boldsymbol{y}}_3|\beta(\hat{\boldsymbol{y}}_2))\mathrm{d}x}{\displaystyle\int_{x'} p(x')P(\hat{\boldsymbol{y}}_1|\alpha(x'))\sum_{\hat{\boldsymbol{y}}_2} P(\hat{\boldsymbol{y}}_2|\alpha(x'))P(\hat{\boldsymbol{y}}_3|\beta(\hat{\boldsymbol{y}}_2))\mathrm{d}x'}. \tag{21}$$

As an alternative receiver for the sawtooth mappings, we will also implement the maximum likelihood (ML) decoder given by

$$\hat{x} = \gamma_{\mathrm{ML}}(y_1,y_3) = \arg\max_x p(y_1,y_3|x). \tag{22}$$

The ML decoder is suboptimal in the sense that it does not minimize the MSE.

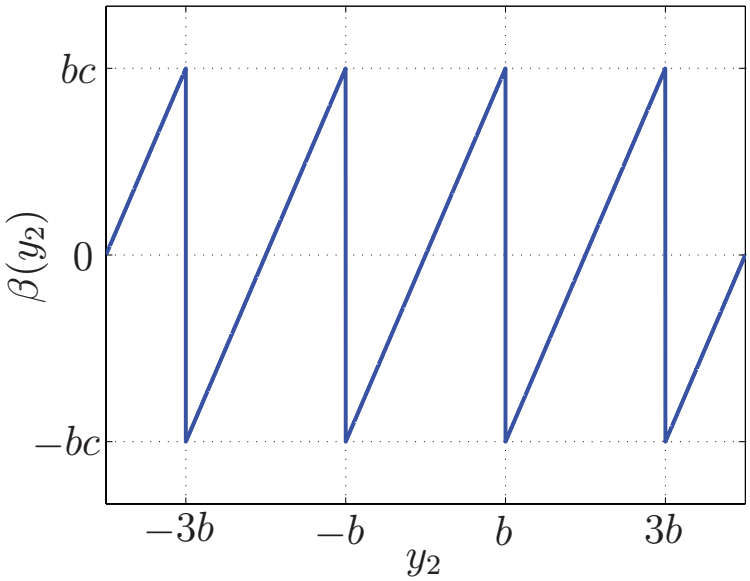


Figure 2: Parameterized sawtooth mapping.

3.4 Design Algorithm

Given the above expressions for the source mapping, the relay mapping, and the receiver it will be possible to optimize the system iteratively. We do this by keeping two parts of the system fixed while we optimize the third part. One common problem with an iterative technique like the one suggested here is that the final solution will depend on the initialization of the algorithm, if the initialization is bad we are likely to end up in a poor local minimum. One method that has proven to be helpful in counteracting this is noisy channel relaxation [9, 16] which works in the following way. A system is first designed for a noisy channel, the solution obtained is then used as an initialization when designing a system for a less noisy channel. The noise is reduced and the process is repeated until the desired noise level is reached. The intuition behind this method is that an optimal system for a noisy channel has a simple structure and is easy to find, as the channel noise is decreased more structure is gradually added to form the final system. Given a scenario where K and L are specified, the design procedure is formally stated in Algorithm 1.

Algorithm 1 Design Algorithm

Require: Initial mappings for β and γ, the channel power gains $\boldsymbol{A} = (a_1, a_2, a_3)$ for which the system should be optimized and the threshold δ that determines when to stop the iterations.
Ensure: Locally optimized α, β, and γ.

Let $\boldsymbol{A'} = (a'_1, a'_2, a'_3)$, where $a'_1 \leq a_1, a'_2 \leq a_2, a'_3 \leq a_3$ (i.e., $\boldsymbol{A'}$ corresponds to a channel which is more noisy than $\boldsymbol{A}$), be the channel power gains for which the system is being optimized.
while $\boldsymbol{A'} \neq \boldsymbol{A}$ **do**
  Increase $\boldsymbol{A'}$ according to some scheme (e.g., linearly).
  Set the iteration index $k = 0$ and $D^{(0)} = \infty$.
  repeat
    Set $k = k + 1$
    Find the optimal source mapping α by using (13).
    Find the optimal receiver γ by using (21).
    Find the optimal relay mapping β by using (18).
    Find the optimal receiver γ by using (21).
    Evaluate the distortion $D^{(k)}$ for the system.
  until $(D^{(k-1)} - D^{(k)})/D^{(k-1)} < \delta$
end while

4 Simulation Results

To evaluate the algorithm we have designed systems for different combinations of K and L. We will compare the performance against some reference systems, given below, and the distortion-rate function for a memoryless Gaussian source [7] using the achievable rate of the compress-and-forward (CF) scheme [17] (assuming orthogonal transmissions).

4.1 Reference Systems

$\boldsymbol{K = L = 1}$**:** For the one-dimensional case we use linear transmission at the source node in conjunction with estimate-and-forward (EF) at the relay as our reference system. For EF, the relay function β is given by $\beta(y_2) = c\mathbf{E}[s_2|y_2]$. It should be noted that in the case of a Gaussian source and linear transmission at the source node, amplify-and-forward is equivalent to estimate-and-forward.

$\boldsymbol{K = 2, L = 1}$**:** In this case, we compare our optimized system with two different reference systems. The first system operates by transmitting the source sample X directly on the channel for both channel uses (scaled to fulfill the power constraint), that is, repetition coding, and uses EF at the relay. This system will be denoted *Linear*. One disadvantage of this scheme

is the repetition coding in the transmission from the source node. To better fill the two-dimensional channel space we propose the following alternative system, denoted *Digital*, where we have taken off-the-shelf components and put them together in a modular fashion. Instead of the source mapping $\alpha(\cdot)$ we use a 16-level Lloyd–Max quantizer [18, 19] followed by a 16-QAM mapping to the channel space. The relay node makes a hard decision on the received signal and modulates the decoded symbol with 16-PAM. At the destination node the received signals are once again decoded with a hard decision and finally x is reconstructed as the expected value of x given the decoded symbols. This system is optimized in the sense that we use a source-optimized quantizer, a good choice of the mapping to QAM symbols (i.e., a mapping that corresponds to a good *index assignment*, so that neighboring quantization levels correspond to neighboring QAM symbols [20]), and an optimal receiver (given the hard decoded received symbols).

$\boldsymbol{K = 1, L = 2:}$ As in the one-dimensional case, we use linear transmission at the source node and study two different relay mappings — a linear repetition code and a digital system, denoted *Linear* and *Digital*, respectively. The Linear relay mapping scales the input to satisfy the power constraint and transmits the same symbol two times. The source symbol, x, is then estimated as the expected value given the received signals. The Digital system performs a 16-level quantization (optimized for the input distribution) and transmits the quantization index using 16-QAM. At the receiver, the quantization index is decoded using a hard decision ML-decoding rule. Finally, the hard decoded index is used in conjunction with the value received on the direct link to find the expected value of the source symbol given these values.

4.2 Implementation Aspects

Before running the design algorithm, β was initialized as a linear mapping and γ was randomly initialized. However, it is important to understand that the use of noisy channel relaxation makes the solution less sensitive to the initialization. In the case of the relay channel, with three different channels, the problem is instead that of choosing a starting point and a path for the noisy channel relaxation. For the case $K = L = 1$, we started at $\boldsymbol{A}_1' = (a_1, -5, -5)$ dB and linearly increased the second and third components one at a time until they reached their corresponding final values. For the other two cases, we started at $\boldsymbol{A}_2' = (-5, -5, -5)$ dB and linearly increased all components simultaneously until they reached $\boldsymbol{A}$. To reduce the complexity of the design algorithm in the case $K = 1$, we fixed α to be a linear scaling (fulfilling the power constraint) followed by a mapping to the closest point in the set $\mathcal{S}$. Lines 7 and 8 were omitted in the design algorithm for these systems. Although there are no proofs that this is the jointly optimal strategy, it can be justified by the fact that linear scaling

is individually optimal for each point-to-point link from the source node in the case $K = 1$. A final note regarding the Lagrange multipliers λ and η. After each iteration in the design algorithm, they were either increased or decreased in small steps depending on whether the used power was too high or too low.

Another important aspect is the number of points, M, in the discretization of the channel space given by (8). In our implementation we have varied M with the total channel SNR, using a lower resolution for low SNRs and a higher resolution for high SNRs. For example, at an SNR of 5 dB we have used $M = 64$ and at 25 dB we have used $M = 512$. Δ has been varied along with M according to $\Delta = 8/(M-1)$; meaning that we have a good approximation of the channel in the interval $[-4, 4]$. There is a tradeoff in the choice of M and Δ, increasing M increases the complexity of the design algorithm. If on the other hand M is too small, the distortion created by the discrete approximation is significant.

Finally, we will give some details of the actual implementation of (13), (18), and (21). All integrals with respect to x and also the source mapping α, have been calculated using a set of training samples which turns the integrals into sums. The size of this set has been 200000 in the case of a $K = 1$ and 10000 in the case of $K = 2$. Since the channel space is approximated by the finite set $\mathcal{S}$, the relay mapping and the receiver can be stored as lookup tables.

4.3 Numerical Results

In the following simulations, we assume that the source mapping and relay mapping are optimized for certain signal-to-noise ratios (SNRs), marked with circles in the figures, but that the receiver has perfect channel state information and therefore adapts to the current channel state using (21). We will mainly study the power efficiency of the relay node, that is, how much power the relay needs to achieve a certain performance. For the one-dimensional case, which we study more extensively, we have also included results showing the power efficiency of the source node for different relay mappings.

$\boldsymbol{K = L = 1}$ **:** If the quality of the link to the relay is better than the direct link, as in Figure 3, the relay can improve the performance significantly. The horizontal power gain[2] of using the optimized system over the linear system is as much as 7–8 dB in the entire region shown. It should be noted that this increase is only due to utilizing the power in a more efficient way and comes at virtually no extra complexity in the relay. The gap to the achievable rate is quite significant, around 6.5–8 dB for the optimized points. This gap will be discussed later on. It is also evident that the op-

[2]We will only consider this gain.

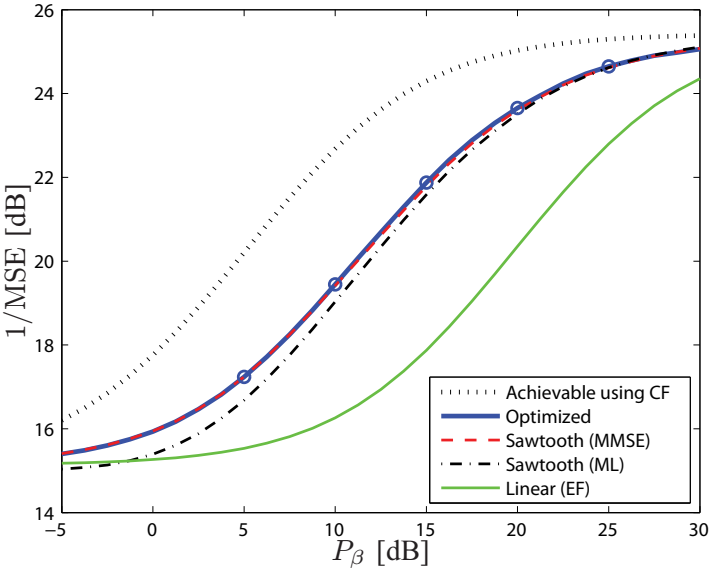


Figure 3: $K = L = 1$ Simulation results when P_β is varied while $P_\alpha = 0$ dB and $a_1 = 15$ dB, $a_2 = 25$ dB, and $a_3 = 0$ dB. The circles mark the points for which the system is optimized.

timized mappings and the sawtooth mappings with MMSE receiver (given by (21)) perform almost the same (the optimized mappings are about 0.1 dB better than the sawtooth mappings at the design points), making them practically impossible to distinguish. It also turns out that the ML detector (given by (22)) performs very close to the optimal MMSE detector, which is encouraging due to its simplicity. It should be emphasized that the sawtooth mappings have been optimized for each SNR point and each detector. A sawtooth mapping which is optimal for the MMSE detector is not necessarily optimal for the ML detector. In Figure 4, we vary the power of the source node. In this case the optimized system manages to follow the achievable curve closely — the gap is only 0.1 dB at $P_\alpha = 5$ dB and increases slightly with the SNR to 0.7 dB at $P_\alpha = 25$ dB. This can be explained as follows, up to some point, say $P_\alpha = 10$ dB, the channel from the relay to the destination is much better than the channels from the source. This implies that all relay mappings perform basically the same as long as they are nondestructive and do not discard any information (c.f. $P_\beta \to$ inf). For this reason, also the linear mapping performs very well. The fact that we are close to the achievable curve strengthens our previous

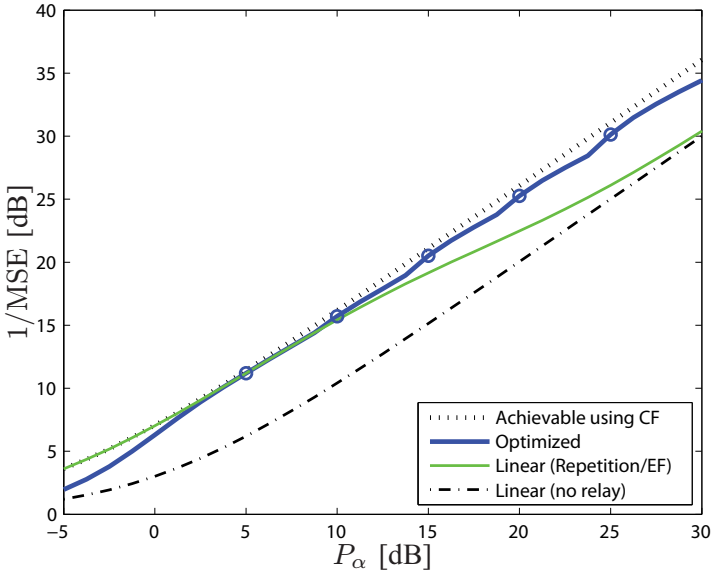


Figure 4: $K = L = 1$ Simulation results when P_α is varied while $P_\beta = 0$ dB and $a_1 = 0$ dB, $a_2 = 5$ dB, and $a_3 = 20$ dB. The circles mark the points for which the system is optimized.

intuitive suggestion that linear transmission at the source node works well for $K = 1$. As the power of the source node increases further, we see that the linear relay mapping approaches the same performance as not using the relay at all. The relatively high noise power on the channel from the relay to the destination makes the information from a linear relay unusable. It is therefore interesting to note how well the optimized mappings follow the achievable curve. As the power of the source node increases, the correlation between $\boldsymbol{y}_1$ and $\boldsymbol{y}_2$ will also increase. The optimized mappings take advantage of this increasing correlation and perform a kind of Wyner–Ziv compression where $\boldsymbol{y}_2$ is used as side information when decoding the information from the relay. An example of how this is done will be given in Section 4.4.

$\boldsymbol{K = 2, L = 1}$ **:** In this case (Figure 5), we have the additional problem of designing a good source mapping, α. The optimized system still has a significant gain over the linear system, ranging from 5 dB at $P_\beta = 5$ dB to 10 dB at $P_\beta = 15$ dB. The digital system performs slightly worse than the linear system. From the figure, the different systems does not seem to reach the same performance as the power of the relay increases. This

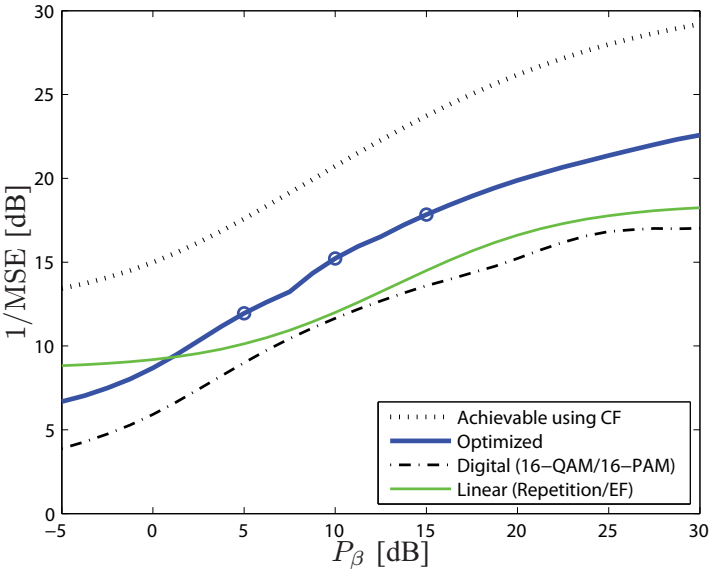


Figure 5: $K = 2, L = 1$ Simulation results when P_β is varied while $P_\alpha = 0$ dB and $a_1 = 5$ dB, $a_2 = 15$ dB, and $a_3 = 0$ dB. The circles mark the points for which the system is optimized.

is in fact true, the achievable curve reaches a limit of 31 dB whereas the performance of the linear system is limited to 18.5 dB. This gap is due to the linear system's inability to produce a two-dimensional distribution that matches the Gaussian channel from the source. The source mapping used in the optimized systems (see Section 4.4) does a better job, but does clearly not achieve the capacity on the two-dimensional channel from the source node either. Similar results for bandwidth expansion curves can be observed in [10].

$\boldsymbol{K = 1, L = 2}$ **:** Changing the situation, having one channel use for the source transmission and two channel uses for the relay transmission, the results are similar to the one-dimensional case as can be seen in Figure 6. The gap to the linear system is around 3 dB and the gap to the achievable curve is ranging from 4.5 dB at $P_\beta = 5$ dB to 8 dB at $P_\beta = 15$ dB.

The significant gap to the achievable curve in most cases can to a large extent be explained by our low-delay one-dimensional approach where we transmit one sample at a time, in contrast to the infinite dimensions used in the proofs for both the distortion-rate function and the achievable rate. An exception to this is when there is no side information available and the

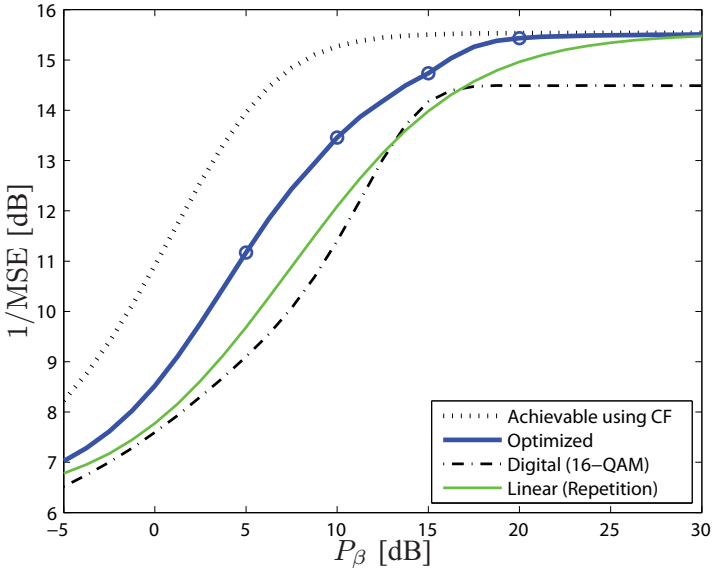


Figure 6: $K = 1, L = 2$ Simulation results when P_β is varied while $P_\alpha = 0$ dB and $a_1 = 5$ dB, $a_2 = 15$ dB, and $a_3 = 0$ dB. The circles mark the points for which the system is optimized.

distribution of the source matches the channel, in which case uncoded transmission is optimal (e.g., transmitting a one-dimensional Gaussian variable on a Gaussian channel).

4.4 Structure of β

$\boldsymbol{K = L = 1}$ **:** Figure 7 shows an example of a typical relay mapping in the one-dimensional case. It is clear that the proposal of sawtooth mappings in Section 3.2 is well motivated. The main reason why this optimized mapping performs better than a linear mapping is the steeper slope, which effectively decreases the impact of the channel noise. Looking at the sawtooth mapping in Figure 2, one could say that decreasing b allows us to increase c (without violating the power constraint) and therefore get lower distortion. It is apparent that the relay mapping is not injective since several input values are mapped to the same output value; this way of reusing output values can be seen as *Wyner–Ziv* compression[3]. The reuse of output values is

[3]The Wyner-Ziv scheme saves *rate* by sending an ambiguous "bin-index" rather than a codeword index. The ambiguity is resolved at the decoder by using the side information to

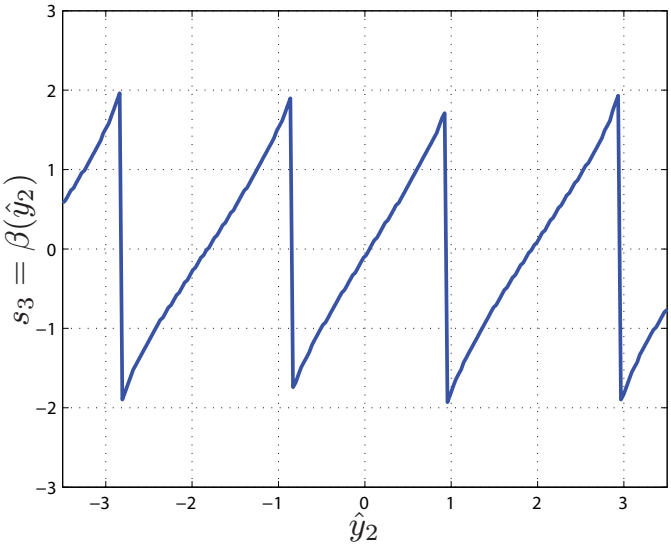


Figure 7: Relay mapping ($K = L = 1$) optimized for $P_\alpha = P_\beta = 0$ dB and $a_1 = 15$ dB, $a_2 = 15$ dB, and $a_3 = 20$ dB.

only possible due to the side information from the direct link. Returning to Figure 2 and assuming that the output of the relay is 0, in this case the side information will provide the necessary information to determine whether y_2 was, for example, $-2b$, 0, or $2b$. However, if b is decreased below a certain threshold (dependent on a_1, a_2, and a_3), the probability of making the wrong decision based on the side information will be significant and the decoder will therefore make large estimation errors. It is in particular the values near the discontinuities that are sensitive to large estimation errors. Looking at the optimized mapping again, one can see that the slope is slightly steeper near the discontinuities. The extra energy spent for these values increases the distance between points in the safe region (far away from the discontinuities) and the critical points (near the discontinuities). This could be the explanation of the slightly better performance of the optimized mappings compared with that of the sawtooth mappings. It is quite remarkable that the design algorithm produces the sawtooth-like mappings despite the fact that the initial relay mapping is *linear*. We believe that this is a consequence of the channel relaxation, especially the fact that a_3 is the last component that is increased, and the Lagrange multipliers and

identify the correct codeword in the bin. In our case, the relay saves *power* by informing the receiver about a set (a "bin") of possible values, rather than a specific value.

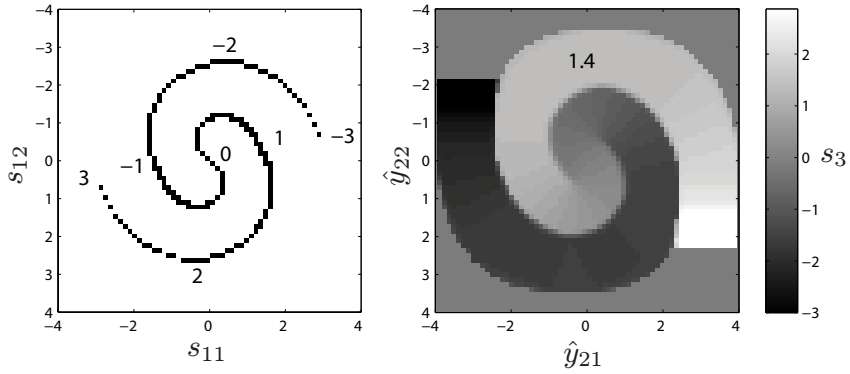


Figure 8: Structure of α (to the left) and β (to the right) ($K = 2, L = 1$) optimized for $P_\alpha = P_\beta = 0$ dB and $a_1 = 5$ dB, $a_2 = 15$ dB, and $a_3 = 10$ dB. In the left part, it is shown how the interval $[3, 3]$ of the one-dimensional input is mapped to the two-dimensional output $\boldsymbol{s}_1 = (s_{11}, s_{12})$. In the right part, the color in the figure together with the colorbar shows how the two-dimensional input, $\hat{\boldsymbol{y}}_2 = (\hat{y}_{21}, \hat{y}_{22})$, is mapped to the one-dimensional output s_3.

how these are updated in small steps.

$\boldsymbol{K = 2, L = 1}$ **:** The source mapping α is now a mapping from the one-dimensional source space to the two-dimensional channel space. An example of such a mapping is shown in the left part of Figure 8, where the curve shows how input symbols in the interval $[-3, 3]$ are mapped to two-dimensional output symbols. The mapping is such that small negative values of x are mapped to one end of the curve and as x is increased the mapping follows the curve to the other end. Values around zero — which are the most likely values for a Gaussian source — are mapped to the center of the curve which lies close to the origin where $\|\boldsymbol{s}_1\|^2$ is small. The transmission power for these values is hence minimized. In contrast, values that are less probable are instead mapped to points in the channel space that use more energy. This structure is due to the Lagrange term in (18); similar results have been been obtained in [9–11]. Due to the high noise level on the direct link, the destination cannot distinguish between different parts of the curve by only looking at the direct link. For example, the receiver will not be able to determine whether 1 or -3 was transmitted since they are mapped to symbols that are close in the channel space. The relay node needs to help the receiver to distinguish which point, or at least which region, of the curve that was transmitted. Looking at the right part of Figure 8, which shows the relay mapping, we can see that this is exactly what the relay does.

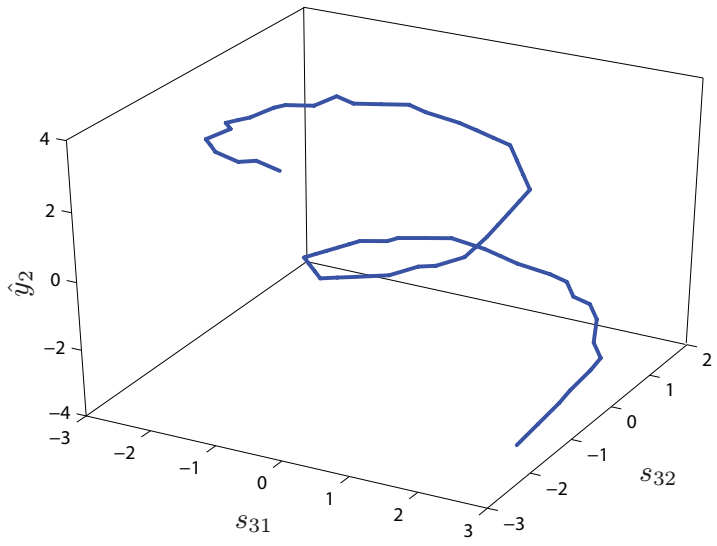


Figure 9: Relay mapping ($K = 1, L = 2$) optimized for $P_\alpha = P_\beta = 0$ dB and $a_1 = 5$ dB, $a_2 = 15$ dB, and $a_3 = 5$ dB. The two-dimensional output, shown on the x- and y-axes, as a function of the one-dimensional input, shown on the z-axis. In other words $\boldsymbol{s}_3 = \beta(\hat{y}_2)$, where $\boldsymbol{s}_3 = (s_{31}, s_{32})$.

Something that is interesting to notice is that the relay is *not* the inverse of the source mapping which it would be if the relay tried to estimate x and send the estimate to the receiver. This is easiest seen by the fact that for some of the outer parts of the curve, the relay uses the same output symbol for large regions (e.g., $s_3 \approx 1.4$ for the upper part of the curve) which means that the relay does not send an estimate of what was received but rather just tells the receiver that the transmitted point was on the upper part of the curve. Using this information the receiver estimates x based on the value received from the direct link conditioned that the transmitted point was on the upper part of the curve.

$\boldsymbol{K = 1, L = 2}$ **:** In Figure 9, we finally show an example of a mapping where the relay performs an expansion — from its one-dimensional input to its two-dimensional output. Once again, there is a reuse of the output symbols which is only possible due to the side information from the direct link. Looking at the spiral from above, a similarity to the polynomial based source–channel codes proposed in [21] can be seen.

5 Conclusions

We have proposed a low-delay scheme for joint source–channel coding over the relay channel. The design also includes optimizing the relay itself. The numerical results show that the joint design works well and gives better performance than the reference systems. We have also provided useful insight into the structure of the (locally) optimized source–channel and relay mappings, and how these mappings together make it possible for the receiver to output a good estimate of the source. The mapping at the relay reuses output symbols and is clearly reminiscent of Wyner–Ziv compression. Based on observing the structure of our optimized systems, we proposed the use of sawtooth mappings for the case of one-dimensional relaying. The sawtooth mappings can in many cases be used instead of the optimized mappings without any performance degradation.

References

[1] E. C. van der Meulen, "Three-terminal communication channels," *Adv. Appl. Prob.*, vol. 3, no. 1, pp. 120–154, 1971.

[2] L. Bao, M. Skoglund, and K. H. Johansson, "Iterative encoder-controller design for feedback control over noisy channels," *IEEE Trans. on Automatic Control*, accepted for publication.

[3] G. N. Nair, F. Fagnani, S. Zampieri, and R. Evans, "Feedback control under data rate constraints: An overview," *Proc. of the IEEE*, vol. 95, no. 1, pp. 108–137, January 2007.

[4] D. Gündüz and E. Erkip, "Source and channel coding for cooperative relaying," *IEEE Trans. on Information Theory*, vol. 53, no. 10, pp. 3454–3475, October 2007.

[5] H. Y. Shutoy, D. Gündüz, E. Erkip, and Y. Wang, "Cooperative source and channel coding for wireless multimedia communications," *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 2, pp. 295–307, August 2007.

[6] T. T. Kim, M. Skoglund, and G. Caire, "On cooperative source transmission with partial rate and power control," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 8, pp. 1408–1418, October 2008.

[7] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. Wiley-Interscience, 2006.

[8] V. Vaishampayan, "Combined source–channel coding for bandlimited waveform channels," Ph.D. dissertation, University of Maryland, 1989.

[9] A. Fuldseth and T. A. Ramstad, "Bandwidth compression for continuous amplitude channels based on vector approximation to a continuous subset of the source signal space," in *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Munich, Germany, April 1997, pp. 3093–3096.

[10] P. A. Floor, T. A. Ramstad, and N. Wernersson, "Power constrained channel optimized vector quantizers used for bandwidth expansion," in *International Symposium on Wireless Communication Systems*, October 2007.

[11] N. Wernersson, J. Karlsson, and M. Skoglund, "Distributed quantization over noisy channels," *IEEE Trans. on Communications*, vol. 57, no. 6, pp. 1693–1700, June 2009.

[12] H. E. III, "Generalized Lagrange multiplier method for solving problems of optimum allocation of resources," *Operations Research*, vol. 11, no. 3, pp. 399–417, 1963.

[13] Y. Shoham and A. Gersho, "Efficient bit allocation for an arbitrary set of quantizers," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol. 9, no. 9, pp. 1445–1453, 1988.

[14] N. Wernersson and M. Skoglund, "Nonlinear coding and estimation for correlated data in wireless sensor networks," *IEEE Trans. on Communications*, vol. 57, no. 10, pp. 2932–2939, October 2009.

[15] S. Yao, M. N. Khormuji, and M. Skoglund, "Sawtooth relaying," *IEEE Communications Letters*, vol. 12, no. 9, September 2008.

[16] S. Gadkari and K. Rose, "Noisy channel relaxation for VQ design," in *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 1996, pp. 2048–2051.

[17] A. Høst-Madsen and J. Zhang, "Capacity bounds and power allocation for wireless relay channels," *IEEE Trans. on Information Theory*, vol. 51, no. 6, pp. 2020–2040, June 2005.

[18] S. P. Lloyd, "Least Squares Quantization in PCM," *IEEE Trans. on Information Theory*, vol. 28, no. 2, pp. 129–137, March 1982.

[19] Y. Linde, A. Buzo, and R. M. Gray, "An algorithm for vector quantizer design," *IEEE Trans. on Communications*, vol. 28, no. 1, pp. 84–95, January 1980.

[20] M. Skoglund, "On channel-constrained vector quantization and index assignment for discrete memoryless channels," *IEEE Trans. on Information Theory*, vol. 45, no. 7, pp. 2615–2622, November 1999.

[21] N. Wernersson, M. Skoglund, and T. Ramstad, "Polynomial based analog source–channel codes," *IEEE Trans. on Communications*, vol. 57, no. 9, pp. 2600–2606, September 2009.

Paper G

Design and Performance of Optimized Relay Mappings

Johannes Karlsson and Mikael Skoglund

Published in *IEEE Transactions on Communications*, September 2010.

Design and Performance of Optimized Relay Mappings

Johannes Karlsson and Mikael Skoglund

Abstract

We look at the three-node relay channel and the transmission of an information symbol from the source node to the destination node. We let the relay be a memoryless function and formulate necessary conditions for the optimality of the relay mapping and the detector. Based on these, we propose a design algorithm to find relay mappings such that the symbol error rate at the destination is minimized. The optimized relay mappings are illustrated for different scenarios and the dependency between the relay mapping and the link qualities is discussed in detail. Furthermore, the performance is compared with existing schemes, such as decode-and-forward, amplify-and-forward, and estimate-and-forward. It is shown that there is a significant gain in terms of decreased symbol error rate if the optimized relay mappings are used.

1 Introduction

Numerous relay strategies have been proposed for the relay channel, with the two most well-known schemes being amplify-and-forward (AF) and decode-and-forward (DF). In this paper, we propose a design of memoryless relay mappings that are optimized[1] for minimum symbol error probability at the destination. Related work of finding optimal memoryless relay mappings includes [1–4]. Reference [1] studies uncoded transmission with BPSK modulation and proves that the probability of error is minimized if the operation at the relay is a Lambert W function. In [2], it is shown that estimate-and-forward maximizes the capacity of the relay channel for a BPSK input

[1]We distinguish between *optimal* which we use to refer to a globally optimal system and *optimized* which potentially is only locally optimal.

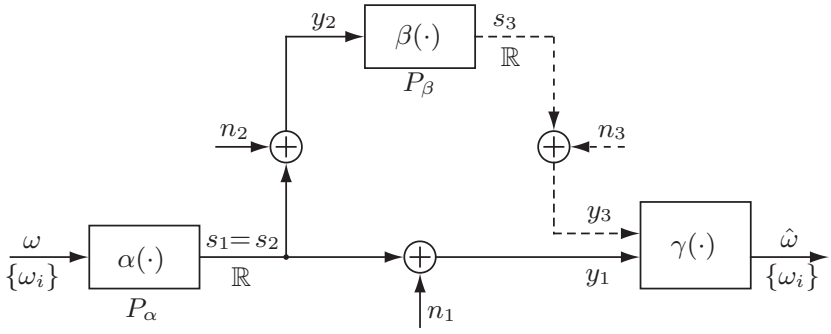


Figure 1: Structure of the system.

signal. The work is carried on in [3], where it is shown that estimate-and-forward also maximizes the signal-to-noise ratio (SNR) at the destination. A related problem, considered in [4], is to find relay mappings that minimize the probability of error for the two-way relay channel. In [4], two nodes communicate with each other via a relay node. During the first phase the two nodes transmit their messages to the relay node over a multiple access channel. In the second phase, the relay transmits to both nodes simultaneously over a broadcast channel. Common to [1–4] is transmission over additive white Gaussian noise (AWGN) channels and that there is no direct link from the source to the destination. If a direct link is available, the operation at the relay can be made more energy efficient by utilizing the side information which is provided by the direct link. One scheme where this is done is constellation-rearrangement [5]. In this scheme the relay first makes a hard decision on the received symbol and then transmits the symbol using a rearranged order of the modulation symbols. Due to the side information, this rearrangement of the symbols is shown to decrease the probability of bit error.

We study uncoded transmission and let the relay be a memoryless mapping such that, at each time instant, it maps its received symbol to an output symbol. In the following sections, we first state the problem more formally and formulate necessary conditions for the optimality of the relay mapping. These conditions are then used to find optimized relay mappings, which are evaluated against the existing schemes DF, AF, EF, and CR in Section 4.

2 Problem Formulation

We consider the three-node relay channel shown in Figure 1. The goal is to transmit an information symbol Ω from the source node to the destination node as reliably as possible. Ω is modelled as an M-ary discrete random variable, uniformly distributed over the set $\{\omega_1, \ldots, \omega_M\}$. At the source node, the function $\alpha : \{\omega_1, \ldots, \omega_M\} \mapsto \mathbb{R}$ modulates an information symbol to a channel symbol $s_1 = \alpha(\omega)$, which is transmitted to the destination node. The transmission is overheard by the relay node which uses the function $\beta : \mathbb{R} \mapsto \mathbb{R}$ to map its observation of the transmission, y_2, to a new channel symbol $s_3 = \beta(y_2)$, which is transmitted to the destination node via a channel that is orthogonal (in time or frequency) to the one used by the source node. We emphasize that we study the transmission of independent uncoded symbols and that the relay is memoryless — meaning that its output only depends on the current input, sometimes referred to as *instantaneous* relaying.

All transmissions are corrupted by additive white Gaussian noise, the received symbols on each channel can therefore be expressed as

$$y_i = s_i + n_i \quad i = 1, 2, 3, \tag{1}$$

where s_i is the transmitted symbol and n_i is independent white Gaussian noise with variance σ_i^2. The "gain" of each link will in the following be expressed in terms of the reciprocal of the noise variance, $a_i = 1/\sigma_i^2$. Since the relay listens to the same channel as the destination we have the equality $s_2 = s_1$. Both the source and the relay are constrained in the sense that they must satisfy an average transmit power constraint

$$\mathbf{E}[S_1^2] \leq P_\alpha, \tag{2}$$

$$\mathbf{E}[S_3^2] \leq P_\beta. \tag{3}$$

At the destination node the received symbols are used to make a decision on the transmitted symbol

$$\hat{\omega} = \gamma(y_1, y_3) \in \{\omega_1, \ldots, \omega_M\}. \tag{4}$$

γ is fully specified by the decision regions $\mathcal{A}_{\omega_i}$ and their complementary regions $\mathcal{A}_{\omega_i}^C$, $i \in \{1, \ldots, M\}$, which are defined as

$$\mathcal{A}_{\omega_i} = \{(y_1, y_3) : \gamma(y_1, y_3) = \omega_i\} \tag{5}$$

$$\mathcal{A}_{\omega_i}^C = \bigcup_{j \neq i} \mathcal{A}_{\omega_j} \tag{6}$$

As performance measure we use the uncoded symbol error rate (SER),

$$P_e = \Pr(\hat{\Omega} \neq \Omega). \tag{7}$$

With "optimal," we therefore refer to a system such that P_e is minimized given the power constraints in (2) and (3). For a fixed modulator, relay mapping, and detector, the whole system can be seen as one equivalent discrete-time channel with certain symbol transition probabilities. The error probability can therefore be further decreased by implementing a powerful channel code, such as an LDPC or turbo code, on top of this effective channel. In this case, a more appropriate design criterion would be to maximize the mutual information between the transmitted and received symbols. However, a simple suboptimal approach is to change our detector, which is optimized for minimum SER, such that it provides soft information instead of hard decisions.

3 Design

With the notation introduced in Section 2 we can write P_e as

$$P_e = \sum_{\omega\in\{\omega_1,\ldots,\omega_M\}} P(\omega) \iint_{(y_1,y_3)\in\mathcal{A}_\omega^C} p(y_1|\alpha(\omega))\times \int_{-\infty}^{\infty} p(y_2|\alpha(\omega))p(y_3|\beta(y_2))\mathrm{d}y_2\,\mathrm{d}y_1\mathrm{d}y_3, \tag{8}$$

where $P(\cdot)$ and $p(\cdot|\cdot)$ denote probability mass functions (pmf:s) and conditional probability density functions (pdf:s), respectively. We let α be represented by any existing modulation scheme, satisfying the power constraint in (2), and now pay attention to finding the relay mapping β and the corresponding detector γ such that P_e is minimized. This design problem is nonconvex and we therefore take the same approach as in [6], where we first discretized the channel space into N equally spaced points according to

$$\mathcal{S} = \Big\{ -\frac{\Delta}{2}(N-1), -\frac{\Delta}{2}(N-3), \ldots, \frac{\Delta}{2}(N-3), \frac{\Delta}{2}(N-1)\Big\}, \tag{9}$$

restricting the output from the relay to satisfy $s_3 \in \mathcal{S}$. All received symbols are then quantized back to $\mathcal{S}$ by the following hard decision decoding rule

$$\hat{y}_i = \arg\min_{y_i'\in\mathcal{S}} |y_i - y_i'| \quad i = 1,2,3, \tag{10}$$

where the hat will be used to indicate that the value has been discretized. This approximation is expected to be good if N is sufficiently large and Δ is small in relation to the standard deviation of the channel noise, σ_i. Next, we formulate necessary conditions for the optimality of β given α and γ and the corresponding necessary conditions on γ given α and β. Having done this, we propose an iterative design algorithm in Section 3.3 for finding an optimized system.

3.1 Optimal Relay Mapping

With the discretized channel space, the optimal relay mapping β, given a fixed α and γ, is given by (see Appendix A)

$$\beta(\hat{y}_2) = \arg\min_{s_3 \in \mathcal{S}} \left(\Pr(\hat{\Omega} \neq \Omega | \hat{y}_2, s_3) + \lambda s_3^2 \right) \tag{11}$$

where

$$\begin{aligned} &\Pr(\hat{\Omega} \neq \Omega | \hat{y}_2, s_3) \\ &= \sum_{\omega \in \{\omega_1, \ldots, \omega_M\}} P(\omega | \hat{y}_2) \sum_{(\hat{y}_1, \hat{y}_3) \in \mathcal{A}_\omega^C} P(\hat{y}_1 | \alpha(\omega)) P(\hat{y}_3 | s_3) \end{aligned} \tag{12}$$

The Lagrange multiplier λ in (11), is used to turn the constrained optimization problem into an unconstrained problem [7,8]. The term λs_3^2 penalizes channel symbols that use high power in favor of channel symbols using low power. $\lambda > 0$ should be chosen such that the power constraint in (3) is fulfilled.

3.2 Optimal Detector

The optimal detector γ, given a fixed α and β, is simply the maximum-a-posteriori (MAP) detector. Consequently, the decision regions are given by

$$\begin{aligned} \mathcal{A}_{\omega_i} = \{(\hat{y}_1, \hat{y}_3) : \ &\Pr(\Omega = \omega_i | \hat{y}_1, \hat{y}_3) \\ &> \Pr(\Omega = \omega_j | \hat{y}_1, \hat{y}_3), \forall j \neq i\}, \end{aligned} \tag{13}$$

where

$$\begin{aligned} &\Pr(\Omega = \omega | \hat{y}_1, \hat{y}_3) \\ &= kP(\Omega = \omega) P(\hat{y}_1, \hat{y}_3 | \Omega = \omega) \\ &= kP(\Omega = \omega) P(\hat{y}_1 | \alpha(\omega)) \sum_{\hat{y}_2 \in \mathcal{S}} P(\hat{y}_2 | \alpha(\omega)) P(\hat{y}_3 | \hat{y}_2), \end{aligned} \tag{14}$$

with k being a constant that is independent of ω.

3.3 Design Algorithm

It is in general hard to optimize β and γ simultaneously since the problem is nonconvex. We therefore propose a design algorithm where we iterate between finding the optimal relay mapping for a fixed detector and vice versa. A common problem with an iterative technique like the one suggested here is that the final solution will depend on the initialization of the algorithm,

if the initialization is bad we are likely to end up in a poor local minimum. One method that has proven to be helpful in counteracting this is noisy channel relaxation [6, 9, 10], which works in the following way. A system is first designed for a noisy channel, the solution obtained is then used as an initialization when designing a system for a less noisy channel. The noise is reduced and the process is repeated until the desired noise level is reached. The intuition behind this method is that an optimal system for a noisy channel has a simple structure and is easy to find, as the channel noise is decreased more structure is gradually added to form the final system. Assuming the source modulation α is given, the design procedure for β and γ is formally stated in Algorithm 1.

Algorithm 1 Design Algorithm

Require: Fixed mapping α as well as an initial mapping for β, the channel power gains $\boldsymbol{A} = (a_1, a_2, a_3)$ for which the system should be optimized and the threshold δ that determines when to stop the iterations.

Ensure: Locally optimized β and γ given α.

Let $\boldsymbol{A'} = (a_1', a_2', a_3')$, where $a_1' \leq a_1, a_2' \leq a_2, a_3' \leq a_3$ (i.e., $\boldsymbol{A'}$ corresponds to a channel which is more noisy than $\boldsymbol{A}$), be the channel power gains for which the system is being optimized.
Find the optimal detector γ by using (13).
while $\boldsymbol{A'} \neq \boldsymbol{A}$ **do**
  Increase $\boldsymbol{A'}$ according to some scheme (e.g., linearly).
  Set the iteration index $k = 0$ and $P_e^{(0)} = 1$.
  repeat
    Set $k = k + 1$
    Find the optimal relay mapping β by using (11).
    Find the optimal detector γ by using (13).
    Evaluate the SER $P_e^{(k)}$ for the system.
  until $(P^{(k-1)} - P^{(k)})/P^{(k-1)} < \delta$
end while

4 Simulation Results

To evaluate the performance of our optimized mappings produced by the design algorithm, the first thing we need to do is to fix M (the size of the information symbol set) and the modulation scheme α. In the following simulations we let $M = 4$ and use pulse amplitude modulation (PAM) at

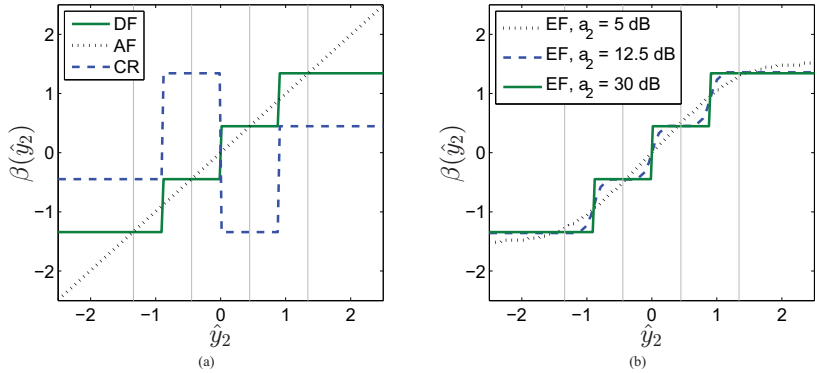


Figure 2: Relay mappings for (a) DF, AF, CR and (b) EF. The vertical lines mark the transmitted PAM points.

the source node, that is,

$$\alpha(\omega_i) = i\Delta_\alpha - \frac{(M+1)}{2}\Delta_\alpha \quad i = 1, \ldots, M, \tag{15}$$

where Δ_α is chosen such that (2) is fulfilled. We repeat that the information symbols are assumed to be uniformly distributed so that all symbols are equally likely.

In the implementation of the design algorithm the following choices were made. As initial mapping for the relay, we used a linear mapping. In line 1, assuming we would like to design a system for $\boldsymbol{A} = (a_1, a_2, a_3)$ dB, we set $\boldsymbol{A'} = (a_1, 5, 5)$ dB. The noisy channel relaxation on line 4 was done in two phases: In the initial phase, for each iteration, the second component of $\boldsymbol{A'}$ was gradually increased until $\boldsymbol{A'} = (a_1, a_2, 5)$ dB. In the second phase, for each iteration, the third component of $\boldsymbol{A'}$ was gradually increased until $\boldsymbol{A'} = \boldsymbol{A}$ dB. It was observed that a better mapping could sometimes be found by increasing $\boldsymbol{A'}$ slightly above $\boldsymbol{A}$ during the noisy channel relaxation and then reducing it to $\boldsymbol{A}$. The channel space was discretized into $N = 256$ different points according to (9) with $\Delta \approx 8/(N-1)$. The discrete approximation allows us to find the relay and detector by an exhaustive search.

4.1 Reference Systems

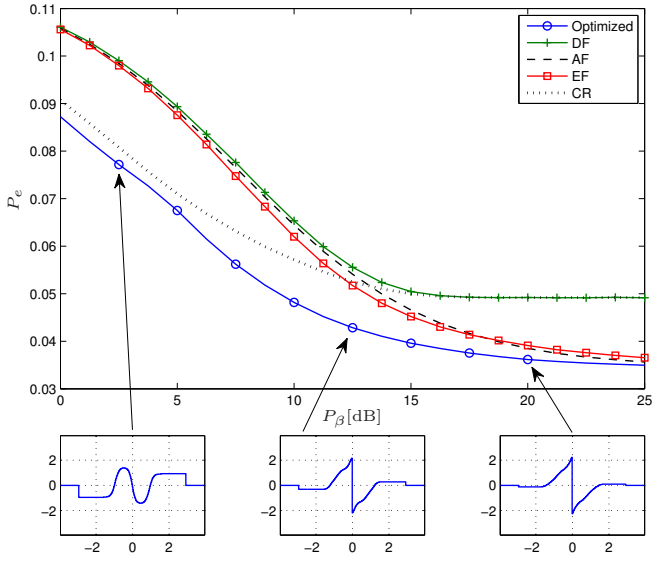
The performance of our optimized mappings will be compared to the following four existing schemes:

- Decode-and-forward (DF) — the relay makes a hard decision on which symbol that was transmitted and transmits the decoded symbol using the same modulation scheme as the source node, see Figure 2(a).

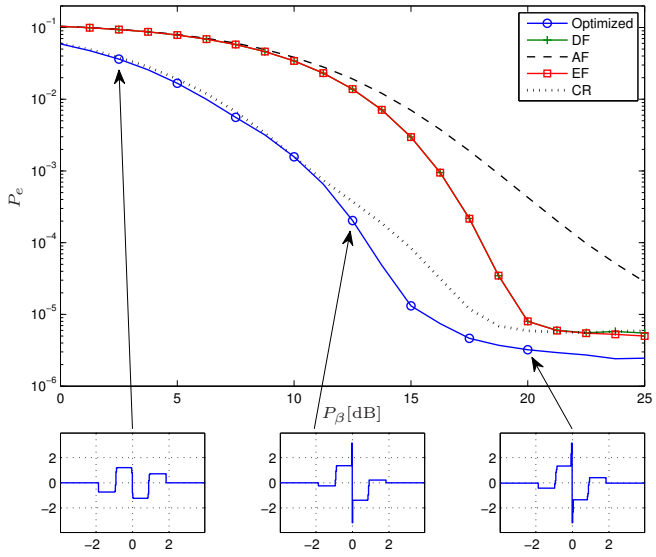
- Amplify-and-forward (AF) — the relay transmits a scaled version of its input, see Figure 2(a).
- Constellation-rearrangement (CR) [5] — similar to DF, the relay makes a hard decision on the received symbol, but instead of using the same modulation scheme as the source the relay uses a rearranged order of the modulation symbols, see Figure 2(a).
- Estimate-and-forward (EF) — the relay transmits $k\mathbf{E}[S_2|y_2]$, where k is a constant such that the power constraint in (3) is satisfied, see Figure 2(b). Something worth noting is that the relay mapping in this case will depend on a_2.

4.2 Numerical Results

The results will be presented for two types of scenarios, in the first scenario (Figure 3) we vary P_β keeping P_α fixed and in the second scenario (Figure 4) we vary P_α keeping P_β fixed. The destination node has perfect channel state information and therefore adapts the detector to the current channel state using (13). Starting with Figure 3(a) with $a_1 = 10$ dB, $a_2 = 10$ dB, and $a_3 = 0$ dB, we can see that the optimized mapping is 3.5–4.5 dB better than EF, which is the best of the conventional methods. CR turns out to work really well in this scenario and closely follows the optimized mapping with a gap of about 0.8–1 dB when $P_\beta < 5$ dB. After this point the CR scheme saturates due to the hard decision inherited from DF. All schemes eventually saturate because of the link to the relay (a_2), which becomes the bottleneck when P_β increases. In Figure 3(b), a_2 has been increased to 20 dB with everything else the same as in the previous case. The gap to EF (and DF) is in this case even bigger — at most 7 dB. In this case the optimized system almost looks identical to the CR scheme, which therefore follows even closer than before. On the other hand, when P_β is increased, the optimized relay mapping is able to utilize the extra power in a more efficient way than CR by providing some soft information on the borders of the PAM points. This will be discussed more in Section 4.3. Moving on to the second scenario, where P_α is varied with $a_1 = a_2 = 0$ dB and $a_3 = 15$ dB in Figure 4(a), we can see that the gain is relatively small in this case — around 0.5–1.5 dB compared to EF and 0.5 dB compared to CR. That all schemes perform almost the same is simply because the role of the relay is minor in this setup since the effective signal-to-noise ratio of the direct link is increasing but the link from the relay to the destination is fixed. However, if we increase a_2 to 5 dB — so that the relay has better knowledge of the transmitted value than the destination — the difference among the schemes becomes more evident as seen in Figure 4(b). EF and DF again have similar shape and performance with a gap to the optimized system of about 1 dB at $P_\alpha = 10$ dB and a gap of 4 dB at $P_\alpha = 14$ dB.

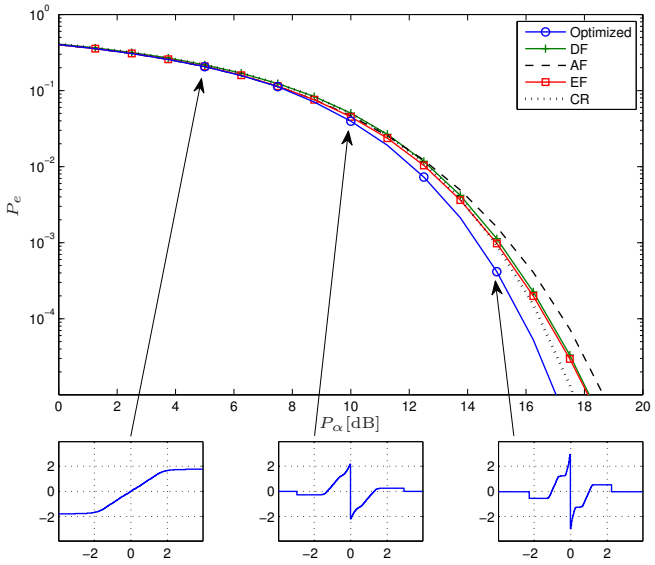


(a) $a_1 = 10$ dB, $a_2 = 10$ dB, and $a_3 = 0$ dB.

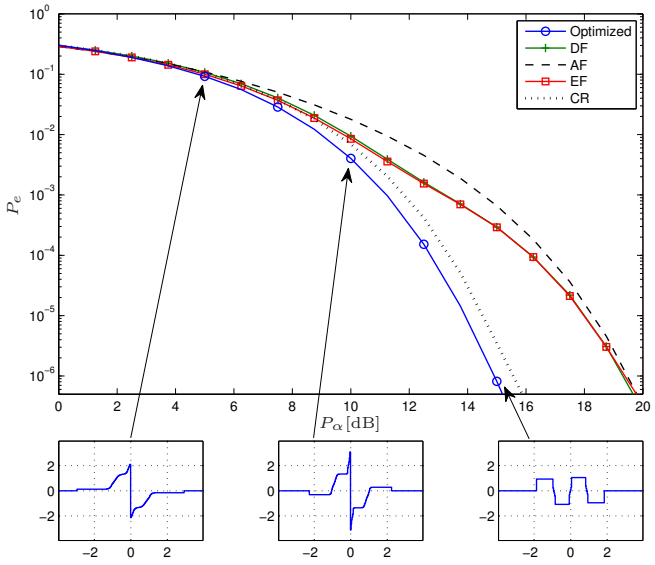


(b) $a_1 = 10$ dB, $a_2 = 20$ dB, and $a_3 = 0$ dB.

Figure 3: Simulation results when the relay power, P_β, is varied while the source power is fixed, $P_\alpha = 1$. The circles mark the SNR points for which the relay mapping has been optimized. A selection of optimized relay mappings are shown below each figure.



(a) $a_1 = 0$ dB, $a_2 = 0$ dB, and $a_3 = 15$ dB.



(b) $a_1 = 0$ dB, $a_2 = 5$ dB, and $a_3 = 15$ dB.

Figure 4: Simulation results when the source power, P_α, is varied while the relay power is fixed, $P_\beta = 1$. The circles mark the SNR points for which the relay mapping has been optimized. A selection of optimized relay mappings are shown below each figure.

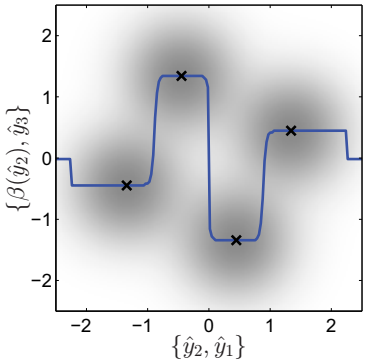


Figure 5: A system designed for $a_1 = 5$ dB, $a_2 = 15$ dB, and $a_3 = 5$ dB with the relay mapping, $\beta(\hat{y}_2)$, superimposed on the joint probability mass function of the received symbols at the destination node, i.e. $P(\hat{y}_1, \hat{y}_3)$. The crosses mark the positions of the received symbols if all links were noiseless.

Again, CR performs quite well with a gap to the optimized system of about 0.6 dB, which is due to the soft information provided by the optimized relay mapping.

4.3 Interpretation and Discussion of β

As seen in the previous section, there is a large gain in terms of decreased P_e by using the optimized relay mapping compared to the otherwise commonly referred AF and DF schemes. In this section, we discuss and explain properties of optimized relay mapping based on our numerical optimizations. In the following discussion, we let $P_\alpha = P_\beta = 1$. One of the most evident characteristics of the optimized relay mappings is that they, in most cases, are not monotonically increasing but instead alternate between positive and negative values. This behavior spreads the symbols in channel space and is due to the side information which the direct link provides. As soon as a_2 reaches a level of approximately 10 dB and above, the non-monotonic mapping is beneficial in comparison to a monotonically increasing mapping — this is true even for as low values of a_1 as 2.5 dB. The explanation to this is that these mappings better fill the channel space at the destination and increase the minimum distance of the transmitted symbols. This can be understood by thinking of the limit as a_2 goes to infinity. In this case, the relay function can be seen as a part of the source node and we can think of it as having two orthogonal channels from the source to the destination, with the constraint that the modulation on the first channel is PAM. If we

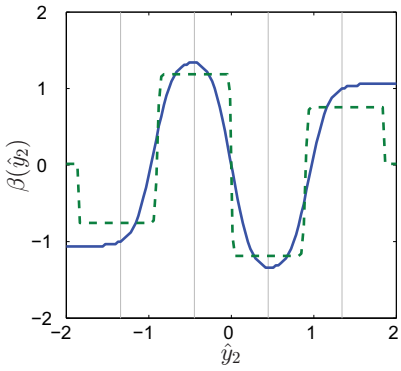


Figure 6: Comparison of optimized relay mappings for a system with $a_1 = 10$ dB, $a_3 = 5$ dB, and $a_2 = 10$ dB (solid) versus $a_2 = 20$ dB (dashed). The vertical lines mark the transmitted PAM points.

assume that $a_1 = a_3$, we would like to use a modulation scheme on the second channel such that the resulting joint modulation is QPSK, which (for a fixed transmit power) maximizes the minimum distance between all transmitted symbols by placing them uniformly spread on a circle. Looking at Figure 5, we notice that this is exactly what the optimized mapping does. The output values from the relay are essentially just a rearrangement of the hard decoded PAM symbols from the source node (c.f. CR).

As a_2 increases for fixed a_1 and a_3, the relay converges into a hard decision device. That is, for low values of a_2 the relay provides soft information of the received symbol (c.f. AF) and for higher values the relay itself makes hard decisions (c.f. DF). However, by using the design algorithm, we can also find relay mappings that work well for intermediate values of a_2 and are not limited to these extremes. The effect of a_2 on the relay mapping is shown in Figure 6.

Another property that is evident appears when a_3 is increased for fixed a_1 and a_2. The relay now starts to transmit soft information that indicates that it is uncertain of its decision at the boundaries of the detection regions. This tells the destination that it should put more trust on the received symbol from the direct link. Observe first in Figure 7, how the relay maps input values close to zero to output values of large amplitude. Looking next at the decision regions, we can see that when the destination receives a value of large amplitude from the relay (i.e. $|\hat{y}_3|$ is large), the border between the two decision regions is close to $\hat{y}_1 = 0$. In other words, the relay tells the destination that the information symbol is either ω_2 or ω_3, but the decision on which of these symbols that was transmitted is left to the direct link.

Finally, we show how the quality of the direct link (i.e. a_1) affects the

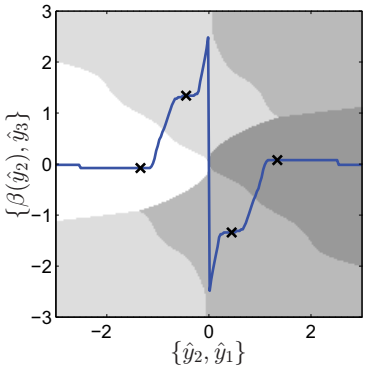


Figure 7: A system designed for $a_1 = 7.5$ dB, $a_2 = 12.5$ dB, and $a_3 = 15$ dB. The relay mapping, $\beta(\hat{y}_2)$, is superimposed on the decision regions at the destination. Note how the relay encodes information about the uncertainty of the information symbol for input values close to zero. This tells the destination to put more trust into the value received from the direct link. The crosses mark the positions, $(\hat{y}_1, \hat{y}_3)$, of the received symbols if all links were noiseless.

relay mapping. If the direct link is weak, the relay tends to make more hard decisions, whereas if the link is strong the amount of soft information from the relay is increased, an example of this is shown in Figure 8. We conclude that the optimal relay mapping should not only depend on the link quality to the relay (c.f. EF), but on all link qualities.

5 Conclusions

We have looked at the three-node relay channel and proposed an algorithm for designing locally optimal relay mappings and the corresponding detector at the destination.

It was shown that the symbol error rate can be significantly decreased by using the optimized relay mapping instead of the otherwise commonly referred schemes AF, DF, or EF. The biggest performance gain stems from the non-monotonic property of the relay mapping, which increases the minimum distance of the transmitted symbols and better fills the channel space at the destination. The proposed system is more flexible than all of the reference systems, since it finds a good tradeoff between soft and hard decisions depending on all link qualities. In the simulations, the optimized systems always perform at least as well as the best of the reference systems and in many cases substantially better.

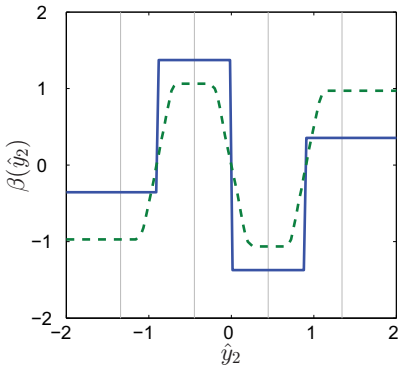


Figure 8: Comparison of optimized relay mappings for a system with $a_2 = 17.5$ dB, $a_3 = 15$ dB, and $a_1 = 5$ dB (solid) versus $a_1 = 15$ dB (dashed). The vertical lines mark the transmitted PAM points.

If used with a channel code, the optimized detector could easily be changed to provide soft information in terms of conditional probabilities instead of hard decisions. One drawback which is unavoidable due to the nature of the problem is that as soon as the quality of one channel changes, the relay mapping and the detector have to be updated. However, the similarity of systems that are optimized for closely related channel qualities suggests that there is some robustness to channel mismatch built in to the system.

Appendix A Optimal Relay Mapping

The optimal relay mapping should minimize P_e given the power constraint in (3). As a first step we turn the constrained optimization problem into an unconstrained problem by using the Lagrange multiplier method [7, 8]

$$\min_{\beta(\cdot):\mathcal{S}\mapsto\mathcal{S}} \left(P_e + \lambda \mathbf{E}[\beta^2(\hat{Y}_2)] \right), \tag{16}$$

where P_e and $\mathbf{E}[\beta^2(\hat{Y}_2)]$ can be written as

$$\begin{aligned} P_e &= \sum_{\hat{y}_2 \in \mathcal{S}} \sum_{\omega \in \{\omega_1,\ldots,\omega_M\}} P(\omega) P(\hat{y}_2|\alpha(\omega)) \times \\ &\quad \sum_{(\hat{y}_1,\hat{y}_3) \in \mathcal{A}_\omega^C} P(\hat{y}_1|\alpha(\omega)) P(\hat{y}_3|\beta(\hat{y}_2)) \\ &= \sum_{\hat{y}_2 \in \mathcal{S}} P(\hat{y}_2) \Big\{ \sum_{\omega \in \{\omega_1,\ldots,\omega_M\}} P(\omega|\hat{y}_2) \times \\ &\quad \sum_{(\hat{y}_1,\hat{y}_3) \in \mathcal{A}_\omega^C} P(\hat{y}_1|\alpha(\omega)) P(\hat{y}_3|\beta(\hat{y}_2)) \Big\} \end{aligned} \tag{17}$$

and

$$\mathbf{E}[\beta^2(\hat{Y}_2)] = \sum_{\hat{y}_2 \in \mathcal{S}} P(\hat{y}_2) \beta^2(\hat{y}_2), \tag{18}$$

respectively. The sums in (17) and (18) have $P(\hat{y}_2)$ as weights and run over all $\hat{y}_2 \in \mathcal{S}$. Since the weights are nonnegative it is clear that the minimization in (16) can be done individually for each relay input, $\hat{y}_2 \in \mathcal{S}$, which gives the expressions in (11) and (12).

References

[1] I. Abou-Faycal and M. Médard, "Optimal uncoded regeneration for binary antipodal signaling," *Proc. of IEEE Int. Conf. on Communications*, vol. 2, pp. 742–746, June 2004.

[2] K. S. Gomadam and S. A. Jafar, "On the capacity of memoryless relay networks," *IEEE International Conference on Communications*, vol. 4, pp. 1580–1585, June 2006.

[3] ——, "Optimal relay functionality for SNR maximization in memoryless relay networks," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 2, pp. 390–401, November 2007.

[4] T. Cui, T. Ho, and J. Kliewer, "Memoryless relay strategies for two-way relay channels: Performance analysis and optimization," *Proc. of IEEE Int. Conf. on Communications*, May 2008.

[5] M. N. Khormuji and E. G. Larsson, "Improving Collaborative Transmit Diversity by Using Constellation Rearrangement," *Proc. of IEEE Wireless Communications and Networking Conf.*, pp. 803–807, March 2007.

[6] J. Karlsson and M. Skoglund, "Optimized low-delay source–channel–relay mappings," *IEEE Trans. on Communications*, no. 5, pp. 1397–1404, May 2010.

[7] H. Everett III, "Generalized Lagrange multiplier method for solving problems of optimum allocation of resources," *Operations Research*, vol. 11, no. 3, pp. 399–417, 1963.

[8] Y. Shoham and A. Gersho, "Efficient Bit allocation for an arbitrary set of quantizers," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol. 9, no. 9, pp. 1445–1453, 1988.

[9] S. Gadkari and K. Rose, "Noisy channel relaxation for VQ design," *Proc. of IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, vol. 4, pp. 2048–2051, May 1996.

[10] A. Fuldseth and T. A. Ramstad, "Bandwidth compression for continuous amplitude channels based on vector approximation to a continuous subset of the source signal space," *Proc. of IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, vol. 4, pp. 3093–3096, April 1997.

Paper H

Closed-Form Sum-MSE Minimization for the Two-User Gaussian MIMO Broadcast Channel

Johannes Kron, Daniel Persson, Mikael Skoglund, and Erik G. Larsson

Submitted to *IEEE Communications Letters*.

Closed-Form Sum-MSE Minimization for the Two-User Gaussian MIMO Broadcast Channel

Johannes Kron, Daniel Persson, Mikael Skoglund,
and Erik G. Larsson

Abstract

We study the Gaussian multiple-input, multiple-output broadcast channel, where a base station with N_T antennas transmits K independent messages to K users, each having a single receive antenna. The messages consist of independent, identically distributed Gaussian random variables and we study linear transmission with an end-to-end distortion criterion. By using an already established uplink/downlink duality and a recently discovered special relation between beamforming vectors and channel vectors, we present a closed-form expression for the optimal power allocation in the two-user case. We also outline an iterative algorithm that finds the optimal power allocation for an arbitrary number of users.

1 Introduction

We consider the Gaussian multiple-input, multiple-output (MIMO) broadcast channel, where a base station with N_T antennas transmits K independent messages to K users, each having a single receive antenna. The messages consist of independent, identically distributed (i.i.d.) complex Gaussian random variables and we study transmission with an end-to-end distortion criterion. The capacity region of the Gaussian MIMO broadcast channel is achievable with dirty-paper precoding (DPC) [1]. Since the messages are independent, we could combine an optimal source code with DPC, achieving the distortion-rate bound for each message, to yield an optimal scheme. However, for the source code and DPC to be optimal, it is in general required that infinite block lengths are used. Motivated by *low-delay* and *low-complexity* constraints, we instead turn to analog transmission using lin-

ear precoding, where the problem is to determine the optimal beamforming vector and power allocation to use for each user's message.

Linear precoding is a well-studied topic. The difficulty in solving the problem lies in the fact that the optimal beamforming vector to each user is dependent both on the power allocation and also on the beamforming vectors used to all other users. Especially interesting to us are the results on signal-to-interference-and-noise ratio (SINR) balancing [2], where an uplink/downlink duality is shown to exist. By using this duality, it is shown that one can determine the optimal beamforming vectors by first solving the much simpler dual uplink problem, which includes uplink power allocation and receive beamforming vectors. The uplink problem is simpler because the beamforming vector for each user is independent of the beamforming vectors for the other users. The optimal downlink solution is obtained by using the dual uplink receive beamforming vectors as transmit beamforming vectors and next finding the optimal downlink power allocation. The distortion criterion we will use is the mean-squared error (MSE), which has a close relation to the SINR. The uplink/downlink duality has been shown to also apply to the MSE region [3].

Algorithms based on convex optimization and iterative techniques for determining optimal power allocation and beamforming vectors were proposed in [2] and [3, 4] for the case of SINR and MSE, respectively. The main contribution of this letter is a closed-form solution for the optimal uplink/downlink MSE power allocation in the case of two users. In the case of more than two users, we outline an iterative algorithm that is of the same complexity order as [4] but conceptually simpler.

2 Problem Formulation

The linear downlink problem is illustrated in Figure 1 and can be formulated as follows: The source variables X_i, $i = 1, \ldots, K$, are to be conveyed to the corresponding ith receiver, where K is the number of users. The encoder multiplies each source variable with a beamforming vector, $\boldsymbol{u}_i \in \mathbb{C}^{N_T}$, and a power scaling variable, $\sqrt{P_i}$, and transmits the sum $\boldsymbol{\alpha}(\boldsymbol{X}) = \sum_{i=1}^{K} \sqrt{P_i}\boldsymbol{u}_i X_i$. The source variables, X_i, are i.i.d. circularly-symmetric complex Gaussian random variables with unit variance and zero mean, that is, $X_i \sim \mathcal{CN}(0, 1)$. The beamforming vectors are of unit norm and the power scaling variables fulfill the relation $\sum P_i \leq P$, where P is the total average power that can be used by the base station. The conjugated channel to user i is described by the vector $\boldsymbol{h}_i \in \mathbb{C}^{N_T}$ such that the received signal can be written as

$$Y_i = \sum_{j=1}^{K} \sqrt{P_j}\boldsymbol{h}_i^H \boldsymbol{u}_j X_j + N_i, \tag{1}$$

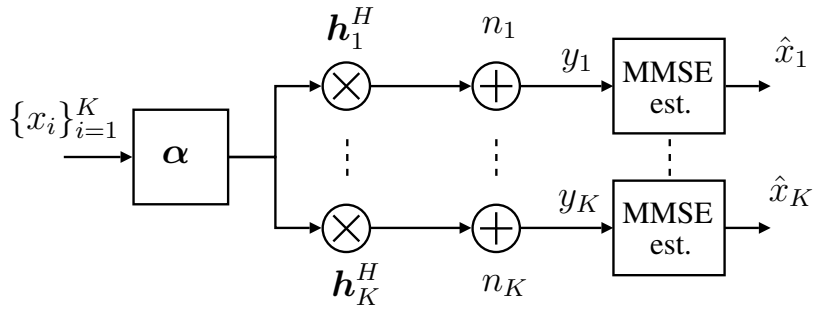


Figure 1: Downlink transmission. The source x_i is transmitted to user i, where the vector $\boldsymbol{h}_i$ models the channel from the base station to the user. The transmission is disturbed by the additive noise n_i.

where $N_i \sim \mathcal{CN}(0,1)$ is additive white Gaussian noise (AWGN). The signal-to-interference-and-noise ratio (SINR) can now be expressed as

$$\mathrm{SINR}_i = \frac{P_i \boldsymbol{h}_i^H \boldsymbol{u}_i \boldsymbol{u}_i^H \boldsymbol{h}_i}{\sum_{j \neq i} P_j \boldsymbol{h}_i^H \boldsymbol{u}_j \boldsymbol{u}_j^H \boldsymbol{h}_i + 1}$$

and the MSE of the linear minimum MSE (MMSE) estimator $\hat{X}_i = \mathbf{E}[X_i|Y_i]$, expressed in terms of the SINR [5], becomes

$$\begin{aligned} \mathrm{MSE}_i = \mathbf{E}[|X_i - \hat{X}_i|^2] &= \frac{\mathbf{E}[|X_i|^2]}{1 + \mathrm{SINR}_i} \\ &= 1 - \frac{P_i \boldsymbol{h}_i^H \boldsymbol{u}_i \boldsymbol{u}_i^H \boldsymbol{h}_i}{\sum_{j=1}^K P_j \boldsymbol{h}_i^H \boldsymbol{u}_j \boldsymbol{u}_j^H \boldsymbol{h}_i + 1}. \end{aligned} \tag{2}$$

We would like to find the jointly optimal beamforming vectors $\{\boldsymbol{u}_i\}$ and power allocation variables $\{P_i\}$ such that the following sum-MSE is minimized

$$\mathrm{MSE} = \sum_{i=1}^K \mathrm{MSE}_i. \tag{3}$$

This is a very hard problem due to the fact that even though the power allocation is fixed, the optimal beamforming vector for user i depends on the beamforming vectors to all other users in a complicated manner, as seen from (2) and (3).

3 Dual Uplink Formulation

In [2], it was shown that the problem of SINR balancing can be solved by first solving a dual uplink problem. In the uplink formulation, each optimal

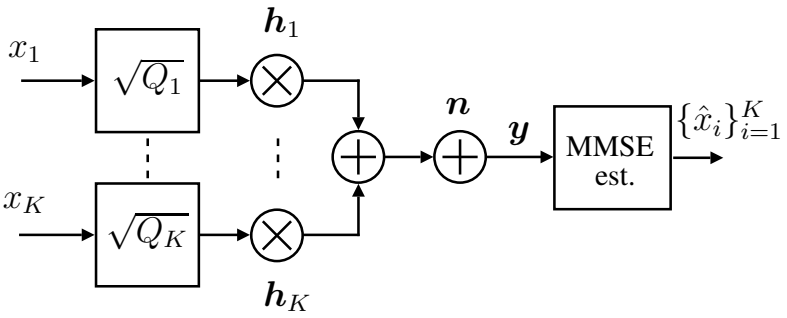


Figure 2: Dual uplink problem. The source variable x_i is transmitted from user i over a MAC, where the vector $\boldsymbol{h}_i$ models the channel from the user to the base station. The transmission is disturbed by the additive noise $\boldsymbol{n}$.

receive beamforming vector is independent of the other beamforming vectors and can therefore easily be found for a given power allocation. Due to the close relationship between SINR and MSE, the duality naturally extends to MSE minimization [3].

The dual uplink problem can be formulated as follows: Each user has a source variable $X_i \sim \mathcal{CN}(0,1)$, $i = 1,\ldots,K$, that is to be conveyed to the base station. At the user node, the source variable is multiplied by a power scaling variable, $\sqrt{Q_i}$, and transmitted to the base station over a multiple-access channel (MAC) as seen in Figure 2. In this formulation, the entries in the channel vector $\boldsymbol{h}_i \in \mathbb{C}^{N_T}$ model the paths from the ith user to each of the N_T antennas at the base station. As before, the transmission is disturbed by an AWGN term $\boldsymbol{N} \sim \mathcal{CN}(0, \boldsymbol{I})$. The received signal at the base station can be expressed as

$$\boldsymbol{Y} = \sum_{i=1}^{K} \sqrt{Q_i}\boldsymbol{h}_i X_i + \boldsymbol{N}, \tag{4}$$

where the power scaling variables should fulfill the constraint $\sum Q_i \leq P$. Each transmitted source variable is now estimated by using the linear beamforming vectors $\boldsymbol{u}_i \in \mathbb{C}^{N_T}$, $\|\boldsymbol{u}_i\| = 1$, and a scalar scaling γ_i such that $\hat{X}_i = \mathbf{E}[X_i|\boldsymbol{Y}] = \gamma_i \boldsymbol{u}_i^H \boldsymbol{Y}$. Thus, the SINR of the ith source variable is

$$\mathrm{SINR}_i = \frac{Q_i \boldsymbol{u}_i^H \boldsymbol{h}_i \boldsymbol{h}_i^H \boldsymbol{u}_i}{\boldsymbol{u}_i^H \Big(\sum_{j \neq i} Q_j \boldsymbol{h}_j \boldsymbol{h}_j^H + \boldsymbol{I} \Big) \boldsymbol{u}_i} \tag{5}$$

and the MSE can be obtained in a similar manner as in (2).

Assuming that an arbitrary power allocation $\boldsymbol{Q} = (Q_1,\ldots,Q_K)$ (i.e., not necessarily fulfilling the power constraint) is given, the uplink formulation

makes it possible to minimize the sum-MSE by individually maximizing each user's SINR. This is done by either of the following two equivalent choices of beamforming vectors [6, Ch. 3]

$$\boldsymbol{u}_i^{\mathrm{opt}}(\boldsymbol{Q}) = \frac{\left(\sum_{j\neq i} Q_j \boldsymbol{h}_j \boldsymbol{h}_j^H + \boldsymbol{I}\right)^{-1} \boldsymbol{h}_i}{\left\|\left(\sum_{j\neq i} Q_j \boldsymbol{h}_j \boldsymbol{h}_j^H + \boldsymbol{I}\right)^{-1} \boldsymbol{h}_i\right\|} \tag{6}$$

$$= \frac{\left(\sum_{j=1}^{K} Q_j \boldsymbol{h}_j \boldsymbol{h}_j^H + \boldsymbol{I}\right)^{-1} \boldsymbol{h}_i}{\left\|\left(\sum_{j=1}^{K} Q_j \boldsymbol{h}_j \boldsymbol{h}_j^H + \boldsymbol{I}\right)^{-1} \boldsymbol{h}_i\right\|}. \tag{7}$$

By inserting the optimal beamforming vector from (6) into (5) and using the relation between SINR and MSE, we obtain

$$\mathrm{MSE}_i^{\mathrm{opt}}(\boldsymbol{Q}) = \frac{1}{1 + Q_i \boldsymbol{h}_i^H \left(\sum_{j\neq i} Q_j \boldsymbol{h}_j \boldsymbol{h}_j^H + \boldsymbol{I}\right)^{-1} \boldsymbol{h}_i}. \tag{8}$$

3.1 Properties of Sum-MSE Minimization

The duality relation is such that an MSE point which is achievable in the uplink can also be achieved in the downlink by using the same beamforming vectors [3]. The uplink/downlink power allocations are in general not equal. Yet, in the special case of *sum-MSE minimization*, it turns out that also the optimal power allocations are equal, that is,

$$P_i^{\mathrm{opt}} = Q_i^{\mathrm{opt}} \qquad i = 1, \ldots, K. \tag{9}$$

This relation stems from the fact that the optimal solution is characterized by the relation

$$\boldsymbol{h}_i^H \boldsymbol{u}_j^{\mathrm{opt}}(\boldsymbol{Q}^{\mathrm{opt}}) = \left(\boldsymbol{h}_j^H \boldsymbol{u}_i^{\mathrm{opt}}(\boldsymbol{Q}^{\mathrm{opt}})\right)^* \qquad \forall i, j. \tag{10}$$

A proof of (9) and (10) was recently presented in [7], where the Karush–Kuhn–Tucker conditions are used to prove the relations. Although (9) and (10) involves $\boldsymbol{Q}^{\mathrm{opt}}$, it is not explicit how to find the actual value of $\boldsymbol{Q}^{\mathrm{opt}}$ without the use of numerical methods. The operational meaning of (10) is that the uplink problem not only is dual to the downlink problem, but at the optimal solutions, the problems are identical.

4 Two-user closed-form solution

We will now use (9) and (10) to derive a closed-form expression for the optimal power allocation in the two-user case. The solution is divided into

three cases depending on whether the channels have the same norm and if they are parallel.

PROPOSITION 1 *Without loss of generality, let* $\|\boldsymbol{h}_1\| \geq \|\boldsymbol{h}_2\| > 0$. *The optimal uplink/downlink power allocation in the two-user sum-MSE case is given by*

$$P_1^{\text{opt}} = \begin{cases} P & \text{if } a_3 = 0, \\ P/2 & \text{if } a_2 = 0, \\ \min\left(\tilde{P}_1, P\right) & \text{otherwise}, \end{cases} \tag{11}$$

where

$$a_1 = 2 + P\|\boldsymbol{h}_2\|^2, \tag{12}$$

$$a_2 = \|\boldsymbol{h}_1\|^2 - \|\boldsymbol{h}_2\|^2 \geq 0, \tag{13}$$

$$a_3 = \|\boldsymbol{h}_1\|^2\|\boldsymbol{h}_2\|^2 - |\boldsymbol{h}_1^H\boldsymbol{h}_2|^2 \geq 0, \tag{14}$$

$$\tilde{P}_1 = \frac{-a_1 + \sqrt{a_1^2 + Pa_1a_2 + a_2^2/a_3}}{a_2}. \tag{15}$$

Proof: Starting with the inner product in (10), we have

$$\begin{aligned} \boldsymbol{h}_2^H\boldsymbol{u}_1^{\text{opt}}(\boldsymbol{P}^{\text{opt}}) &\overset{(a)}{=} \frac{\boldsymbol{h}_2^H\left(P_2^{\text{opt}}\boldsymbol{h}_2\boldsymbol{h}_2^H + \boldsymbol{I}\right)^{-1}\boldsymbol{h}_1}{\left\|\left(P_2^{\text{opt}}\boldsymbol{h}_2\boldsymbol{h}_2^H + \boldsymbol{I}\right)^{-1}\boldsymbol{h}_1\right\|} \\ &\overset{(b)}{=} \frac{\boldsymbol{h}_2^H\left(\boldsymbol{I} - \frac{P_2^{\text{opt}}\boldsymbol{h}_2\boldsymbol{h}_1^H}{1+P_2^{\text{opt}}\boldsymbol{h}_2^H\boldsymbol{h}_2}\right)\boldsymbol{h}_1}{\left\|\left(\boldsymbol{I} - \frac{P_2^{\text{opt}}\boldsymbol{h}_2\boldsymbol{h}_2^H}{1+P_2^{\text{opt}}\boldsymbol{h}_2^H\boldsymbol{h}_2}\right)\boldsymbol{h}_1\right\|} \\ &\overset{(c)}{=} \frac{\boldsymbol{h}_2^H\left((1+P_2^{\text{opt}}\boldsymbol{h}_2^H\boldsymbol{h}_2)\boldsymbol{I} - P_2^{\text{opt}}\boldsymbol{h}_2\boldsymbol{h}_2^H\right)\boldsymbol{h}_1}{\left\|\left((1+P_2^{\text{opt}}\boldsymbol{h}_2^H\boldsymbol{h}_2)\boldsymbol{I} - P_2^{\text{opt}}\boldsymbol{h}_2\boldsymbol{h}_2^H\right)\boldsymbol{h}_1\right\|} \\ &\overset{(d)}{=} \frac{\boldsymbol{h}_2^H\boldsymbol{h}_1}{\left\|\left((1+P_2^{\text{opt}}\boldsymbol{h}_2^H\boldsymbol{h}_2)\boldsymbol{I} - P_2^{\text{opt}}\boldsymbol{h}_2\boldsymbol{h}_2^H\right)\boldsymbol{h}_1\right\|} \qquad (16) \\ &\overset{(e)}{=} \frac{(\boldsymbol{h}_1^H\boldsymbol{h}_2)^*}{\left\|\left((1+P_1^{\text{opt}}\boldsymbol{h}_1^H\boldsymbol{h}_1)\boldsymbol{I} - P_1^{\text{opt}}\boldsymbol{h}_1\boldsymbol{h}_1^H\right)\boldsymbol{h}_2\right\|}, \qquad (17) \end{aligned}$$

where (a) follows from the definition of $\boldsymbol{u}_1^{\text{opt}}$ in (6), (b) from using the matrix inversion lemma, (c)–(d) are basic manipulations, and (e) follows from the symmetry in (10). Since the numerators in (16) and (17) are equal we must

have that also the denominators are equal. By using this equality, squaring both sides and using the relation $P = P_1^{\text{opt}} + P_2^{\text{opt}}$, we get a second-order equation, which when solved gives the expression in (15). Since the solution to a second-order equation in general is unbounded, we need $\min(\tilde{P}_1, P)$ to make sure that the power usage does not exceed the power limit. In the final expression in (11), we also have to consider the cases when a_2 or a_3 are equal to zero. $a_3 = 0$ corresponds to the channels being parallel, in which case the optimal (linear) strategy is to allocate all power to the user with the strongest channel. $a_2 = 0$ on the other hand corresponds to the case where the two channels have the same norm, in which case the power should be divided equally among the users. P_1^{opt} is continuous in the special cases mentioned above and can also be found by taking the limit of $\tilde{P}_1$ as a_2 or a_3 approaches zero. If we have $a_2 = a_3 = 0$, it can be shown by inserting (2) into (3) that the sum-MSE is invariant to the power allocation and only dependent on the sum. ∎

The power allocation for the second user is easily found since $P = P_1^{\text{opt}} + P_2^{\text{opt}}$. Once the optimal power allocation has been determined, the optimal beamforming vectors can be calculated using (6) or (7). For users that are assigned zero power, the beamforming vectors can be arbitrarily chosen. The relation in (10) is therefore not necessarily fulfilled for these users.

5 $K > 2$ Users

A generalization to $K > 2$ users is not straightforward and finding a closed-form solution for this case is still an open problem. We briefly outline an efficient iterative algorithm that takes advantage of the relation in (10). By inserting (7) into (10) and defining $\boldsymbol{A} \triangleq \sum_{i=1}^{K} P_i \boldsymbol{h}_i \boldsymbol{h}_i^H + \boldsymbol{I}$, it can be shown that, for the optimal power allocation, $C_j = \|\boldsymbol{A}^{-1}\boldsymbol{h}_j\| = C$ is constant for all j. Given an initial power allocation (e.g., uniform), we propose an algorithm that is based on evaluating C_j for all users and calculating the arithmetic mean $\bar{C}$ for users with positive power allocation. Next, the power allocation P_j is increased proportionally to $C_j - \bar{C}$ for all users. By updating the powers iteratively, we have been able to find the optimal power allocation for systems with hundreds of users. In each step, one has to take care so that no user has negative power, in which case the power is set to zero, and that the power constraint is fulfilled. The complexity of the algorithm is of the same order as the iterative algorithm presented in [4], that is, $\mathcal{O}\big(L(N_T^3 + KN_T^2)\big)$, where L is the number of iterations. In comparison, algorithms based on convex optimization have a complexity of $\mathcal{O}\big(K^{6.5}N_T^{6.5}\big)$ [3].

6 Conclusions

We have considered sum-MSE minimization for the Gaussian MIMO broadcast channel. By using recently discovered properties of this problem, we have derived a closed-form expression for the optimal power allocation in the two-user scenario and proposed a conceptually simple and efficient algorithm that handles an arbitrary number of users. It is not clear how or whether the closed-form solution can be generalized to more than two users. Our hope is that the solution we present can inspire future research on the more general case.

References

[1] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Trans. on Information Theory*, vol. 52, no. 9, pp. 3936–3964, September 2006.

[2] M. Schubert and H. Boche, "Solution of the multiuser downlink beamforming problem with individual SINR constraints," *IEEE Trans. on Vehicular Technology*, vol. 53, no. 1, pp. 18–28, January 2004.

[3] S. Shi, M. Schubert, and H. Boche, "Downlink MMSE transceiver optimization for multiuser MIMO systems: duality and sum-MSE minimization," *IEEE Trans. on Signal Processing*, vol. 55, no. 11, pp. 5436–5446, November 2007.

[4] A. Mezghani, M. Joham, R. Hunger, and W. Utschick, "Transceiver design for multi-user MIMO systems," in *International ITG Workshop on Smart Antennas*, 2006.

[5] P. Viswanath, V. Anantharam, and D. N. C. Tse, "Optimal sequences, power control, and user capacity of synchronous CDMA systems with linear MMSE multiuser receivers," *IEEE Trans. on Information Theory*, vol. 45, no. 6, pp. 1968–1983, 1999.

[6] R. A. Monzingo and T. W. Miller, *Introduction to adaptive arrays*. SciTech Publishing, 2004.

[7] A. J. Tenenbaum and R. S. Adve, "Minimizing sum-MSE implies identical downlink and dual uplink power allocations," *arXiv:0912.3323v2 [cs.IT]*, 2010, to appear in IEEE Communications Letters.