Power and Bandwidth Efficient Coded Modulation for Linear Gaussian Channels

Niek J. Bouman*
Centrum Wiskunde & Informatica
Amsterdam, The Netherlands
bouman@cwi.nl

Harm S. Cronie*†
Ecole Polytechnique Fédérale de Lausanne
Switzerland
harm.cronie@epfl.ch

Abstract

A scheme for power- and bandwidth-efficient communication on the linear Gaussian channel is proposed. A scenario is assumed in which the channel is stationary in time and the channel characteristics are known at the transmitter. Using interleaving, the linear Gaussian channel with its intersymbol interference is decomposed into a set of memoryless subchannels. Each subchannel is further decomposed into parallel binary memoryless channels, to enable the use of binary codes. Code bits from these parallel binary channels are mapped to higher-order near-Gaussian distributed constellation symbols. At the receiver, the code bits are detected and decoded in a multistage fashion. The scheme is demonstrated on a simple instance of the linear Gaussian channel. Simulations show that the scheme achieves reliable communication at 1.2 dB away from the Shannon capacity using a moderate number of subchannels.

1 Introduction

We consider the classical problem of efficient and reliable communication over the continuous-time linear Gaussian channel [4]. Despite of its age, this channel model is still often used, mainly because of its simplicity and practical relevance. Our objective is to develop a block coded modulation scheme that is capable of achieving power- and bandwidth-efficient communication in the high-SNR regime for an acceptable complexity. We assume that the channel is stationary in time and that the impulse response of the channel is known at the transmitter. We focus on channel instances for which the capacity-achieving band [4] is a single frequency interval, such that capacity can be achieved using serial (single-carrier) transmission. We do not consider multi-carrier transmission.

We deal with the intersymbol interference (ISI) which is due to the linear filter in the channel model by decomposing the channel through interleaving into a number of memoryless subchannels, similar to [10, 8]. Consequently, conventional error-correcting block codes (for memoryless channels) can be applied. Multilevel coding [7] enables the use of binary error-correcting codes combined with spectral-efficient modulation. In this paper we employ state-of-the-art optimized binary low-density parity-check (LDPC) codes [5, 9]. To achieve capacity on the memoryless subchannels, the subchannel inputs should be Gaussian distributed. Therefore, we use *superposition modulation*, with which we can generate near-Gaussian channel inputs [3, 2]. This type of modulation achieves a shape gain [4] over equiprobable signaling with ordinary pulse amplitude modulation (PAM) constellations. At the side of the receiver, multistage

*Most of the work has been done while both authors were affiliated to the Signals & Systems Group, Universiteit Twente.

†Harm Cronie is supported by A. Shokrollahi's Grant 228021-ECCSciEng of the European Research Council

detection and decoding with hard decision feedback is employed. As in [10], an *a posteriori* probability (APP) detector is used that is based on the Bahl-Cocke-Jelinek-Raviv (BCJR) algorithm.

The novel contribution of our paper is the application and evaluation of superposition modulation [2] to the linear Gaussian channel, or stated alternatively, the extension of the binary communication system proposed in [10] to higher-order and spectral-efficient modulation.

The remaining part of the paper is organized as follows. Section 2 considers the channel model and the characteristics of the capacity-achieving input process. We take the continuous-time channel description as a starting point, and we discuss a method to convert it into an equivalent discrete-time model. In Section 3, we decompose the nonbinary ISI channel into binary memoryless subchannels, we consider suitable mappings from bits to constellation symbols and we discuss the method of detection and decoding. Also, we deal with the determination of achievable rates. Section 4 demonstrates the scheme on a simple instance of the linear Gaussian channel.

2 Channel Model

The linear Gaussian channel is defined by

$$y(t) = h(t) * x(t) + n(t),$$

in which $x(t)$ and $y(t)$ are respectively the continuous-time input and output signal, $h(t)$ is the continuous-time impulse response and $n(t)$ is a realization of a zero-mean Gaussian noise process with power spectral density $N(f)$. The asterisk denotes linear convolution. The capacity of this channel is achieved using an input signal that has a zero-mean Gaussian amplitude distribution and the optimal water-pouring power spectral density [4].

At the transmitter, the continuous-time input signal $x(t)$ can be constructed by modulating a train of pulse shapes with a discrete-time information sequence $\{X_i\}$, i.e., $x(t) = \sum_i X_i h_T(t - iT)$. To achieve capacity, the information sequence $\{X_i\}$ should consist of independent zero-mean Gaussian random variables. The symbol response $h_T(t)$ is chosen such that it shapes the flat input spectrum to the optimal water-pouring power spectral density.

At the receiver, as is known from detection theory, the continuous-time received signal may again be discretized, without loss of optimality, using a matched filter and a sampler. By assuming a whitened matched filter (WMF), the entire continuous-time part of the communication system may be abstracted as a digital finite impulse response (FIR) filter with additive white Gaussian noise (AWGN)

$$Y_n = \sum_{k=0}^{\nu} h_k X_{n-k} + W_n, \qquad W_n \sim \mathcal{N}(0, \sigma^2) \tag{1}$$

in which ν denotes the length of the ISI causing tail, or simply the *memory length.* As described in [4], a causal filter with a minimum-phase response can be derived by performing a spectral factorization. We apply this methodology for two examples in Section 4, because a discrete-time channel representation is required for various parts of the simulations as well as for the APP detectors.

3 Multilevel Coding and Multistage Decoding

Figure 1 shows the proposed block-coded modulation scheme. The following subsections discuss the key concepts in detail.

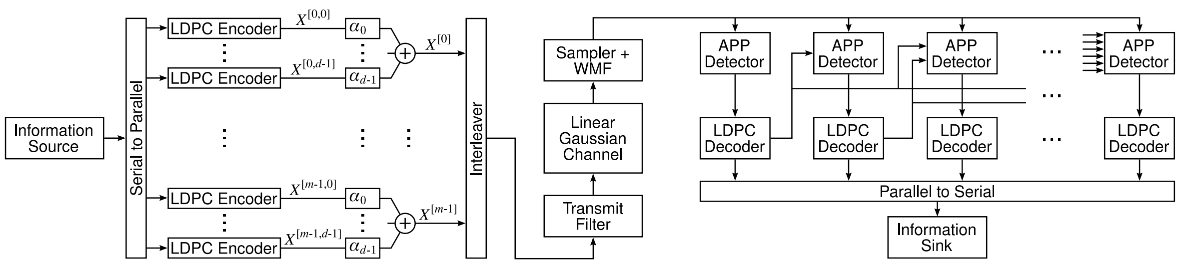

Figure 1: The proposed block coded modulation scheme, employing LDPC channel coding, superposition modulation, interleaving and multistage detection and decoding.

3.1 Dealing with the Intersymbol Interference by Interleaving

Consider a length-mN sequence of output symbols $\{Y_n\}_{n=0}^{mN-1}$, obtained by passing independent inputs $\{X_n\}$ through the channel defined in (1). We reshape this sequence into a m-by-N matrix:

$$\begin{bmatrix} Y_0 & Y_m & \dots & Y_{(N-1)m} \\ Y_1 & Y_{m+1} & \dots & Y_{(N-1)m+1} \\ \vdots & \vdots & & \vdots \\ Y_{m-1} & Y_{2m-1} & \dots & Y_{Nm-1} \end{bmatrix}. \tag{2}$$

If $m > \nu$, the elements of an arbitrary row of (2) are independent, and can be viewed as outputs of a memoryless channel. This implies that with sufficiently deep interleaving, we can decompose the intersymbol interference channel of (1) into a set of m memoryless *subchannels*. The nth output of the ith subchannel is given by

$$Y_n^{[i]} = X_n^{[i]} + N_n^{[i]}. \tag{3}$$

The input $X_n^{[i]}$ and output $Y_n^{[i]}$ of the memoryless channel correspond respectively to the input and output of (1) as X_{nm+i} and Y_{nm+i}. The additive noise $N_n^{[i]}$ is composed of the noise random variable W and ν inputs to other subchannels

$$N_n^{[i]} = W_{nm+i} + \sum_{k=1}^{\nu} h_k X_{nm+i-k}. \tag{4}$$

Note that we have essentially only rewritten (1) into (3) and (4). It follows from (4) that if the X_i are Gaussian distributed, all subchannels have AWGN.

Under assumption of ideally coded subchannels and multistage decoding with decision feedback, and for a particular input density f_X, the sum of the constrained capacities of the memoryless subchannels converges to the constrained capacity of the original ISI channel for $m \to \infty$, as proved in [10, 8]. Remember that for our channel of interest, the *un*constrained capacity can be achieved when f_X is a zero-mean Gaussian density.

3.2 Multilevel Modulation, Constellations and Shaping

In this section we consider how to create a near-Gaussian distributed subchannel input $X^{[i]}$. We assume that channel encoders are present that emit bits that are approximately i.i.d. For this reason, we apply multilevel coding [7] and decompose each memoryless subchannel introduced in Section 3.1 into d parallel *binary* memoryless subchannels. Let us denote the nth input of the jth binary subchannel that belongs to the ith nonbinary memoryless subchannel as $X_n^{[i,j]}$. A possible way to obtain a suitable higher-order discrete signal set is to add the d binary subchannel inputs in the field of the real numbers [3, 2], i.e.

$$X_n^{[i]} = \sum_{j=0}^{d-1} X_n^{[i,j]}, \quad \text{for} \quad X_n^{[i,j]} \in \{-1, +1\}. \tag{5}$$

In the limit for $d \to \infty$, the discrete distribution of $X_n^{[i]}$ converges to the continuous Gaussian distribution by the central limit theorem. In [2], constellations generated by (5) are termed *binomial constellations.* Alternatively, the bits may be scaled prior to addition with positive weights α_j [2];

$$X^{[i]} = \sum_{j=0}^{d-1} \alpha_j X^{[i,j]}, \quad \text{where} \quad X^{[i,j]} \in \{-1, +1\}. \tag{6}$$

The weights can be found by numerical optimization of the mutual information between input and output of a memoryless AWGN channel. The resulting *numerically optimized constellations* work especially well in the high-SNR regime.

3.3 Detection and Decoding

Let us consider the chain rule of mutual information [6], that is adapted to the particular subchannel structure of the proposed scheme

$$\mathrm{I}(\boldsymbol{X}^{mN}; \boldsymbol{Y}^{mN}) = \sum_{i=0}^{m-1} \sum_{j=0}^{d-1} \mathrm{I}\left(\boldsymbol{X}^{[i,j],N}; \boldsymbol{Y}^{mN} \middle| \Psi^{[i,j]}\right) \tag{7}$$

in which $\boldsymbol{X}^{[i,j],N}$ is the jth length-N binary codeword that contributes to symbols belonging to the ith interleave, and $\Psi^{[i,j]}$ represents the set of previously decoded (and error-free) codewords, comprising the hard side information $\Psi^{[i,j]} = \{\boldsymbol{X}^{[a,b],N} | \, da + b < di + j\}$, $a, b \in \mathbb{Z}$. It follows from (7) that information from previous decodings should be used to detect a certain subchannel, except for the first binary subchannel in which $\Psi^{[0,0]} = \emptyset$. Hence, the receiver of the proposed scheme recovers the information bits using multistage detection and decoding.

The *a posteriori* probabilities for the code bits are efficiently computed using the BCJR algorithm. The side information is incorporated in the BCJR algorithm's decisions by altering the transition probabilities in the trellis. The APP detector outputs log-APP ratios,

$$L_n^{[i,j]} = \ln \frac{\Pr(X_n^{[i,j]} | \boldsymbol{Y}^{mN})}{1 - \Pr(X_n^{[i,j]} | \boldsymbol{Y}^{mN})}$$

that are provided to the decoder. To limit the complexity, the BCJR algorithm is executed once per binary subchannel; the detector and decoder do not iteratively exchange information.

3.4 Achievable Rates

The (unconstrained) capacity of the continuous-time linear Gaussian channel can be exactly computed using the water pouring capacity formulas. However, for the same channel no solutions currently exist for the exact calculation of the achievable rate that belongs to a particular discrete signal constellation and input distribution, i.e. the constrained capacity. To determine this rate, we use the lower and upper bound presented in [1]. The bounds are based on the asymptotic equipartition property [6] and are estimated with a Monte Carlo method. The lower and upper bound are given by

$$\hat{\mathrm{I}}(X;Y)_{\text{low.bound}} = \hat{\mathrm{H}}(Y) - \hat{\mathrm{H}}(Y|X), \quad \text{and} \quad \hat{\mathrm{I}}(X;Y)_{\text{upp.bound}} = \hat{\mathrm{H}}(Y) - \mathrm{H}(W).$$

All $\mathrm{H}(\cdot)$ denote differential entropies. $\hat{\mathrm{H}}(Y)$ is computed as $\hat{\mathrm{H}}(Y) = -\frac{1}{N}\log_2 \Pr(\boldsymbol{y}^N)$, in which $\boldsymbol{y}^N$ represents a length-N vector (N very large) of simulated channel output, obtained by passing a length-N vector $\boldsymbol{x}^N$ consisting of superposition-modulated i.i.d. bits through the channel model defined in (1). $\Pr(\boldsymbol{y}^N)$ is estimated using the forward pass of the BCJR algorithm, which computes metrics based on the following channel law

$$\Pr(Y_n|\boldsymbol{X}_{n-\varphi}^n) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \frac{-\left(Y_n - \sum_{k=0}^{\varphi} h_k X_{n-k}\right)^2}{2\sigma^2} \tag{8}$$

in which $\boldsymbol{X}_{n-\varphi}^n$ denotes the vector $[X_{n-\varphi}, X_{n-\varphi+1}, \ldots, X_n]$, σ^2 is the noise variance of the equivalent discrete-time channel model that was introduced in (1) and φ is an integer in the range $0 \le \varphi \le \nu$. This 'truncation parameter' φ controls the trade-off between the tightness of the bounds and the computational complexity of the simulation. The conditional differential entropy $\hat{\mathrm{H}}(Y|X)$ (used in the lower bound) is estimated as

$$\hat{\mathrm{H}}(Y|X) = -\frac{1}{N}\sum_{n=0}^{N-1} \log_2 \Pr\left(y_n|\boldsymbol{x}_{n-\varphi}^n\right).$$

The term $\mathrm{H}(W)$ in the upper bound denotes the differential entropy of the Gaussian noise, which is known in closed form [6], i.e. $\mathrm{H}(W) = \frac{1}{2}\log_2 2\pi e\sigma^2$.

The achievable rate on a binary subchannel from the proposed system can be directly estimated from the output of the APP detector belonging to that binary subchannel [10], $R^{[i,j]} = \mathrm{E}\left[1 - \log_2\left(1 + \exp\left(-X^{[i,j]}L^{[i,j]}\right)\right)\right]$. The APP detectors are based on (8), which means that φ also controls the trade-off between achievable subchannel rates and the complexity of the detectors. The achievable rate on the firstly detected binary subchannel ($R_{[0,0]}$) is independent of m but does depend on the mapping from bits to constellation symbols. The order of detection and decoding and the ordering of the scaling factors in (6) (for constellation types other than binomial) influence the distribution of the subchannel rates. Because it is harder to design good binary codes for very low or very high rates, these orders can be altered to obtain moderate subchannel rates.

3.5 LDPC Component Codes

The proposed coded modulation method results in a set of binary memoryless channels for which binary codes can be used. In this paper we use binary LDPC codes [5]. LDPC codes are amenable to analysis for binary symmetric channels. Furthermore, the structure of the codes can be optimized to lead to a near-capacity performance. In this paper we omit details regarding to the actual optimization and refer to [9] for more details.

4 Example: First-order RC Low-Pass Channel

We demonstrate the proposed scheme by considering baseband transmission of real-valued symbols over a simple instance of the Gaussian channel: an electrical resistor-capacitor low-pass filter circuit with additive Gaussian noise. The Fourier-domain transfer function of the filter is given by

$$H(f) = \frac{1}{1 + \mathrm{j}2\pi\tau f}. \tag{9}$$

in which f denotes the frequency and τ is the time constant, equal to the product of the resistance and capacitance. The cut-off frequency of this filter lies at $f = (2\pi\tau)^{-1}$. For simplicity, we assume a flat noise power spectral density, i.e. $N(f) = N_0/2$.

We consider a scenario in which there is only a power (SNR) constraint. The width of the capacity-achieving band depends on the SNR. Using water pouring [4], the spectral efficiency (i.e., the capacity per dimension, expressed in bits/dim) can be derived in closed form,

$$\frac{C}{2W} = \frac{1}{\ln 2} - \frac{\arctan\left(\sqrt{3\frac{E_s}{N_0}}\right)}{\sqrt{3\frac{E_s}{N_0}}\ln 2} \text{ [bits/dim]}, \tag{10}$$

where C is the capacity (in bits/s), W is the one-sided bandwidth (in Hz), E_s the energy per symbol and N_0 the one-sided noise power spectral density. The spectral efficiency is plotted in Figure 2(a). Note that $C/2W$ approaches a limit for infinite SNR

$$\lim_{E_s/N_0 \to \infty} \frac{C}{2W}\left(\frac{E_s}{N_0}\right) = \frac{1}{\ln 2} \approx 1.44 \text{ [bits/dim]}. \tag{11}$$

We target at a rate of 1 bit/dim. The capacity curve crosses this rate at $E_s/N_0 \approx 8.1$ dB. For this SNR, we compute the equivalent discrete-time channel representation [4]. The result is shown in Table 1. In this example, we use a two-bit binomial constellation. This constellation consists of equispaced non-equiprobable signal points, see also Figure 2(b). To limit the computational complexity of the achievable-rate simulations and of the detector of the example system, we use a five-coefficient channel model in the BCJR algorithm, i.e. $\varphi = 4$. With these settings, simulations indicate that the achievable rate, for $m \to \infty$, is lower bounded by 0.982 bits/dim and upper bounded by 0.992 bits/dim. We choose $m = 3$, resulting in a system comprising $d \times m = 2 \times 3 = 6$ binary subchannels. From the estimated subchannel rates for this system (which can be found in Table 2), we conclude that a rate of 0.978 bits/dim is achievable. Based on the rates printed in Table 2, we have designed six LDPC component codes of blocklength 10^5. We have simulated all codes independently, by assuming perfect side information in each level. Also, we have simulated the entire system with all codes collaborating together, to incorporate the possibility of error propagation into the simulation. The bit-error rate (BER) versus SNR curves of the individual codes and of the entire system are plotted in Figure 3. The system's mean rate (computed as the sum of all component code rates divided by m) amounts to 0.966 bits/dim. This rate equals the capacity at $E_s/N_0 \approx 7.3$ dB. From the overall performance curve, we find that the system operates reliably at 8.5 dB. Hence, the gap to capacity is 1.2 dB.

Table 1: Discrete-Time Representation of the RC Low-Pass Channel

E_s/N_0	σ^2	h_0	h_1	h_2	h_3	h_4	...
8.1 dB	0.485	1	1.114	0.456	0.269	0.103	...

Table 2: Achievable Rates for the RC Low-Pass Channel Example

Binary-Subchannel Rates	0.284	0.351	0.416
	0.463	0.626	0.792
Cumulative Rates	0.738	0.978	1.208
Mean Rate		0.978	

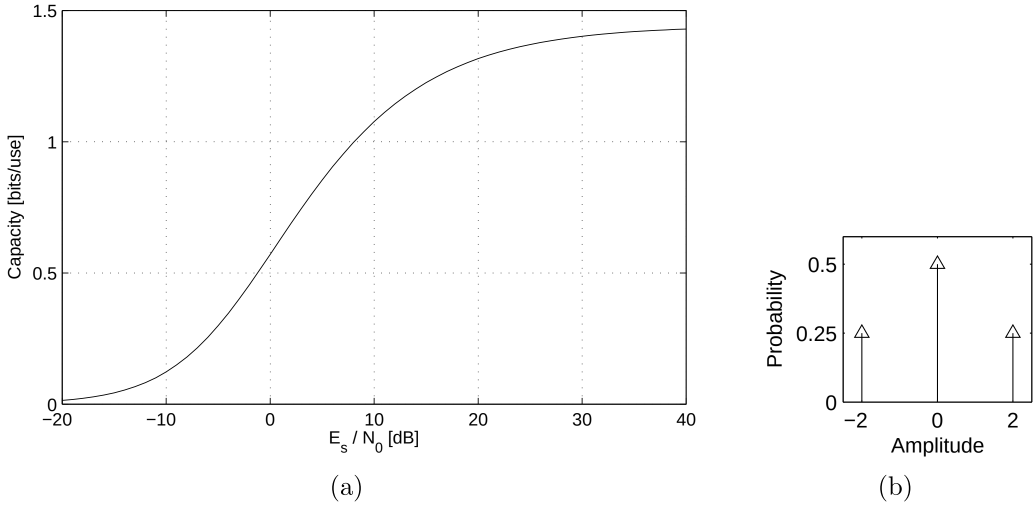

Figure 2: (a) Spectral efficiency curve of the low-pass filter channel. It approaches a limit of 1.44 bit/dim for infinite SNR. (b) Two-bit binomial signal constellation; non-equiprobable equispaced signal points.

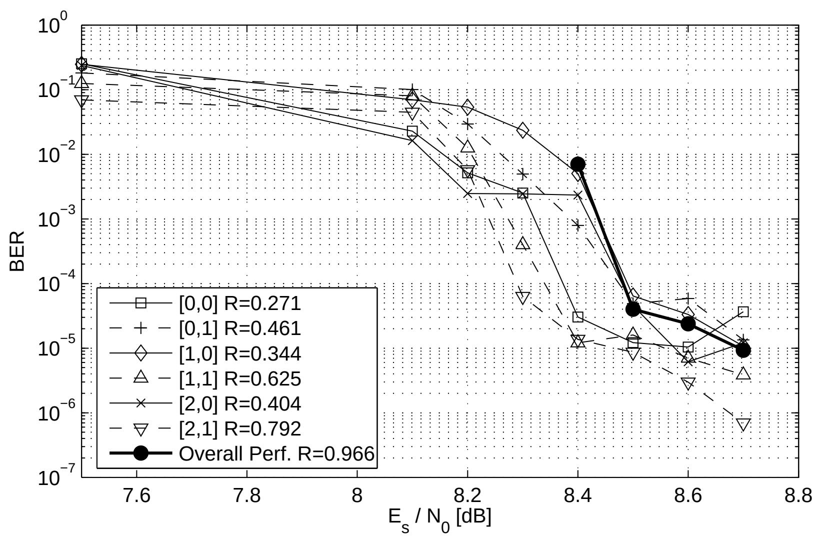

Figure 3: Bit-error rate simulations for the power constraint example.

5 Conclusion

We considered the problem of spectral-efficient communication over the linear Gaussian channel. We have proposed a coded modulation scheme that combines interleaving (to deal with the intersymbol interference), spectral-efficient modulation and multistage decoding. We discussed methods to find achievable information rates. To demonstrate the scheme, we considered an example on a simple instance of the linear Gaussian channel. With the use of LDPC component codes, the scheme achieves a low bit-error rate in this example at 1.2 dB away from the Shannon capacity.

References

[1] D. Arnold, H.-A. Loeliger, and P. O. Vontobel. Computation of information rates from finite-state source/channel models. In *Proc. 40th Annual Allerton Conference on Communication, Control and Computing*, 2002.

[2] H. S. Cronie. Power and bandwidth efficient communication on the AWGN channel by the superposition of binary error-correcting codes. 2007. submitted to IEEE Trans. Inf. Theory.

[3] Long Duan, Bixio Rimoldi, and Rüdiger Urbanke. Approaching the AWGN channel capacity without active shaping. In *Proc. IEEE Int. Symp. Inf. Theory*, page 374, 1997.

[4] G. D. Forney and G. Ungerboeck. Modulation and coding for linear gaussian channels. *IEEE Trans. Inf. Theory*, 44(6), October 1998.

[5] R. G. Gallager. *Low-Density Parity-Check Codes.* PhD thesis, Cambridge, MA: MIT Press, 1963.

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

[7] H. Imai and S. Hirakawa. A new multilevel coding method using error correcting codes. *IEEE Trans. Inf. Theory*, 23:371–377, May 1977.

[8] T. Li and O. M. Collins. A successive decoding strategy for channels with memory. *IEEE Trans. Inf. Theory*, 53(2):628–646, Feb 2007.

[9] T. J. Richardson, R. L. Urbanke, and M. A. Shokrollahi. Design of capacity-approaching irregular low-density parity-check codes. *IEEE Trans. Inf. Theory*, 47:619–637, Feb 2001.

[10] J. B. Soriaga, H. D. Pfister, and P. H. Siegel. Determining and approaching achievable rates of binary intersymbol interference channels using multistage decoding. *IEEE Trans. Inf. Theory*, 53(4):1416–1429, April 2007.