

Trellis Precoding for the Broadcast Channel

Wei Yu and John M. Cioffi

Electrical Engineering Department
Stanford University, Stanford, CA 94305, U.S.A.
e-mails: {weiyu, cioffi}@dsl.stanford.edu

Abstract— This paper considers the vector Gaussian broadcast channel where a single transmitter with multiple antennas sends independent information to multiple receivers. An achievable rate region is derived by decomposing the broadcast channel into a series of single-user channels with non-causal side information. The side information may be completely pre-subtracted using precoding techniques. A practical trellis precoding method is presented. Trellis precoding can be viewed as a generalization of the Tomlinson-Harashima precoder. By taking into account the entire non-causal side-information sequence, a trellis precoder gives an additional shaping gain up to 1.53dB compared to a Tomlinson precoder.

I. INTRODUCTION

Consider a typical multiuser communications scenario with a single base station and many users in geographically different locations. The uplink direction, where multiple transmitters send independent information to a single receiver, is often modeled as a multiple access channel. The downlink direction, where a single transmitter sends independent information to multiple receivers is often modeled as a broadcast channel. The multiple access direction has been well-studied in recent years. The capacity region for the multiple access channel is well-known, and the capacity can be achieved using a single-user code at each transmitter and a multiuser joint detector at the receiver. The broadcast situation, on the other hand, is much less understood. This is mainly because the capacity region for the broadcast channel is still not known, except in the special degraded case. This paper focuses on the broadcast problem. We will present two results for the non-degraded Gaussian vector broadcast channel with multiple antennas at the transmitter. First, we will show that an achievable rate region for the vector broadcast channel can be derived by decomposing the broadcast channel into a series of single-user channels with non-causal side information. Secondly, we will show that a practical trellis precoding method based on a generalization of Tomlinson-Harashima precoder can be used to approach the capacity for channels with side information, thus approaching the broadcast rate region as well. These results will shed some light both on the computation of the achievable rate region for the broadcast channel, and on practical means to achieve it.

This work was supported in part by a Stanford Graduate Fellowship, and in part by Alcatel, Fujitsu, Samsung, France Telecom, IBM, Voyan, Sony, and Telcordia.

II. BROADCAST CHANNEL

In a broadcast channel, a single transmitter attempts to send independent information to two or more receivers at the same time. The capacity region for a two-receiver broadcast channel refers to the set of rate-pairs (R_1, R_2) that can be achieved simultaneously with arbitrarily small probability of error. A complete characterization of the capacity region for the broadcast channel is not yet known except in the special degraded case, where one user's received signal is a corrupted version of the other user's received signal. The Gaussian broadcast channel with a single antenna at the transmitter is an example of a degraded broadcast channel. Consider

$$\begin{aligned} y_1 &= x + n_1 \\ y_2 &= x + n_2, \end{aligned} \quad (1)$$

where x is the input signal subject to a power constraint P , y_1, y_2 are the outputs, n_1 and n_2 are the additive Gaussian white noises with variances σ_1^2 and σ_2^2 , respectively. If $\sigma_1 < \sigma_2$, then y_2 is a degraded version of y_1 . The capacity region for a degraded broadcast channel is achieved using interference subtraction at the receiver [1]. The idea is to divide the total power into $P_1 = \alpha P$ for the first user and $P_2 = (1 - \alpha)P$ for the second user ($0 \leq \alpha \leq 1$). Then, the following rate pair is achievable:

$$R_1 = \frac{1}{2} \log \left(1 + \frac{P_1}{\sigma_1^2} \right) \quad (2)$$

$$R_2 = \frac{1}{2} \log \left(1 + \frac{P_2}{\sigma_2^2 + P_1} \right). \quad (3)$$

Because y_2 is a degraded version of y_1 , information transmitted to y_2 can be decoded by y_1 without error. Thus, y_1 can subtract the interference due to P_2 , and in effect y_1 has a cleaner channel with noise σ_1^2 instead of $\sigma_1^2 + P_2$. It can be shown [2] that this interference subtraction scheme is in fact optimum for the degraded broadcast channel.

The study of the broadcast channel is related to the study of the channel with transmitter side information. Consider the following channel,

$$y_k = x_k + s_k + n_k, \quad (4)$$

where x_k and y_k are the transmitted and the received signals respectively, and s_k is an interfering signal that is known to the transmitter but not to the receiver. In

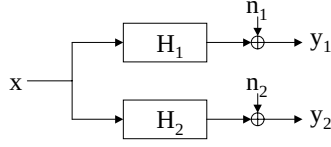


Fig. 1. The vector broadcast channel

this case, the optimum coding scheme turns out to involve the pre-subtraction of s_k at the transmitter. In fact, in a surprisingly result known as “writing-on-dirty-paper”, Costa [3] showed that if both s_k and n_k are i.i.d. Gaussian, and if the entire non-causal realization of s_k is known to the transmitter prior to transmission, the capacity of the channel with interference s_k is the same as if s_k does not exist. In addition, the optimal transmit signal x_k is statistically independent of s_k .

The “dirty-paper” result gives us another way to derive the Gaussian broadcast channel capacity. Suppose that two codebooks $\mathcal{C}_1$ and $\mathcal{C}_2$ with average power P_1 and P_2 are used to send messages to the two receivers y_1 and y_2 , respectively. In each coding block, the transmitter sends out $c_1 + c_2$, where $c_1 \in \mathcal{C}_1$ and $c_2 \in \mathcal{C}_2$ are codewords intended for y_1 and y_2 . If the two codebooks are designed independently, then each receiver sees the other user’s signal as noise. But, the transmitter knows both codewords in advance. So, $\mathcal{C}_1$ can be designed to pre-subtract the codeword to be emitted from $\mathcal{C}_2$. Effectively, x to y_1 is a channel with non-causal side information available at the transmitter but not at the receiver. Thus, Costa’s “dirty-paper” result can be used to achieve a transmission rate to y_1 that is as high as if the interference is not present. Now, the channel to y_2 still sees the first user’s signal as noise. This is because after pre-subtraction, the signal for y_1 is statistically independent of the signal for y_2 . Thus, the rate pair (R_1, R_2) in (2) and (3) is achievable where R_1 does not see P_2 as interference, and R_2 sees both P_1 and σ_2^2 as noise. This gives an alternative derivation for the degraded Gaussian broadcast channel capacity.

Curiously, this derivation does not use the fact that y_2 is a degraded version of y_1 . In fact, y_1 and y_2 may be interchanged and the following rate pair is also achievable:

$$R_1 = \frac{1}{2} \log \left(1 + \frac{P_1}{\sigma_1^2 + P_2} \right) \quad (5)$$

$$R_2 = \frac{1}{2} \log \left(1 + \frac{P_2}{\sigma_2^2} \right). \quad (6)$$

However, when $\sigma_1 < \sigma_2$, the above rate region is strictly smaller than the degraded rate region. Thus, the order of interference pre-subtraction matters.

The idea of subtracting interference at the transmitter instead of subtracting interference at the receiver is attractive because it is also applicable to the non-degraded broadcast channels. Consider the case where the trans-

mitter and receivers are equipped with multiple antennas as shown in figure 1:

$$\begin{aligned} \mathbf{y}_1 &= \mathbf{H}_1 \mathbf{x} + \mathbf{n}_1 \\ \mathbf{y}_2 &= \mathbf{H}_2 \mathbf{x} + \mathbf{n}_2, \end{aligned} \quad (7)$$

where $\mathbf{x}$ is the transmitted vector signal in $\mathbf{R}^n$, $\mathbf{y}_1$ and $\mathbf{y}_2$ are received vector signals in $\mathbf{R}^m$, $\mathbf{H}_1$ and $\mathbf{H}_2$ are channel matrices of dimension $m \times n$, and $\mathbf{n}_1$ and $\mathbf{n}_2$ are Gaussian random vectors with covariance matrices $\mathbf{Z}_1$ and $\mathbf{Z}_2$ respectively. In general, the two channels are not degraded version of each other. Also, because the channel matrices $\mathbf{H}_1$ and $\mathbf{H}_2$ may not have the same eigenvectors, they cannot be decomposed into parallel independent degraded broadcast channels in general. (An important exception is the ISI channel which can be decomposed by a discrete Fourier transform [4].) Nevertheless, the “dirty-paper” result may be generalized to the vector case to implement interference pre-subtraction at the transmitter.

Lemma 1: Consider the channel $\mathbf{y}_k = \mathbf{x}_k + \mathbf{s}_k + \mathbf{n}_k$, where $\mathbf{s}_k$ and $\mathbf{n}_k$ are independent i.i.d. vector Gaussian processes. Suppose that non-causal knowledge of $\mathbf{s}_k$ is available at the transmitter but not at the receiver. The capacity of the channel is the same as if $\mathbf{s}_k$ is not present. Further, the capacity-achieving $\mathbf{x}_k$ is statistically independent of $\mathbf{s}_k$.

This result has been noted by several authors [5] [6]. A simple proof due to Lapidoth [7] goes as follows. The vector Gaussian channel may be decomposed into parallel scalar channels along the eigenvectors of the noise covariance. The resulting parallel Gaussian scalar channels have independent noise, but have correlated interference. However, since the interference is completely known at the transmitter, the “writing-on-dirty-paper” result can be applied individually on each of the scalar channels regardless of the correlation among the interference. This achieves the vector channel capacity as if the interference does not exist. This lemma directly leads to the following achievability result for the vector broadcast channel.

Theorem 1: For the vector Gaussian broadcast channel (7) under a power constraint P , the following rate pairs are achievable:

$$R_1 = \frac{1}{2} \log \frac{|H_1 S_1 H_1^T + H_1 S_2 H_1^T + Z_1|}{|H_1 S_2 H_1^T + Z_1|} \quad (8)$$

$$R_2 = \frac{1}{2} \log \frac{|H_2 S_2 H_2^T + Z_2|}{|Z_2|}, \quad (9)$$

where $|\cdot|$ denotes the matrix determinant, S_1 and S_2 are positive semidefinite matrices satisfying the power constraint: $\text{trace}(S_1 + S_2) \leq P$.

In the above statement, the choice that S_1 is to be the pre-subtracted signal is arbitrary. It is possible to reverse the order and pre-subtract S_2 first, or in fact split the

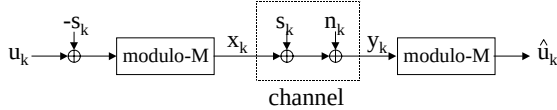


Fig. 2. Tomlinson-Harashima precoder

transmitting signal into more than two users, and achieving the rate-splitting points. The entire achievable rate region is therefore the convex hull of all achievable rate pairs satisfying the power constraint. It is worth noting that the above does not represent a new achievability result for the broadcast channel. Rather, it is an instance of Marton's region [8]. It is not known whether this achievability region is the actual capacity region. In fact, it is not even clear whether the union of all such achievable rate pairs is convex, so time-sharing may be necessary.

Nevertheless, the above result provides a new coding technique for the broadcast channel that is superior to a single-user design that ignores the interference. The first such example was demonstrated in the context of DSL [9], where a broadcast channel with a transmitter with N antennas, and N receivers with a single antenna each, is considered. The channel is modeled as:

$$\mathbf{y} = H\mathbf{x} + \mathbf{n} \quad (10)$$

where $\mathbf{y}$ is an $N \times 1$ vector with each component as a receiver. The choice for S_i , $i = 1, \dots, N$ is made as follows. Let $H = RQ$ be a QR-decomposition, where Q is orthogonal, and R is triangular. The transmit direction for each user is chosen to be the row vectors of Q . More precisely, let $\mathbf{x} = Q^T \mathbf{u}$, where $\mathbf{u}$ is a vector with each component as a data stream for each receiver. Then, the broadcast channel is decomposed into N parallel channels. Because R is triangular, the first channel is a usual Gaussian channel; the second channel has non-causal side information from the first channel; and the third channel has non-causal side information from the first two channels, etc. Thus, an interference pre-subtracting scheme for each of the parallel sub-channels can be used to eliminate interference completely.

In fact, this QR-type decomposition was independently considered by Caire and Shamai [10], who also proved that QR is rate-sum optimal in both low and high SNR regions, and a modified version is optimal for all SNR's in two-user channels. In a recent result [11], the achievable rate pair in Theorem 1 is shown to be indeed the optimum rate-sum point for arbitrary number of users with arbitrary number of transmit and receive antennas. Therefore, the "writing-on-dirty-paper" type of interference pre-subtraction has a wide range of applicability. This motivates us to find practical precoding methods to implement interference pre-subtraction. Since the precoding of vector channels can be implemented by precod-

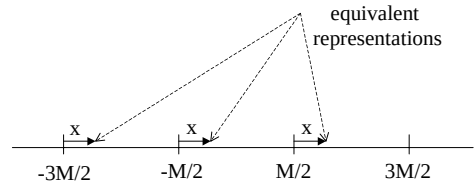


Fig. 3. Modulo Arithmetic

ing a series of scalar channels, the rest of the paper will concentrate on the precoding of a Gaussian scalar channel with non-causal side information.

III. PRECODING

A. Tomlinson-Harashima Precoder

The pre-subtraction of multiuser interference is very much like the pre-subtraction of the intersymbol interference using a Tomlinson-Harashima precoder [12] [13] as shown in figure 2. Let s_k be the total ISI at time k . The transmitter knows s_k at each time instant, so in order to convey an intended symbol u_k , a precoder sends $x_k = u_k - s_k$ to compensate for the interference. However, s_k may be large, so x_k may exceed the power constraint. The trick is to constrain the intended symbol u_k to lie within $[-M/2, M/2)$. Instead of sending x_k , the encoder sends x_k modulo- M . Likewise, the decoder also does a modulo- M operation. The modulo- M operation reduces the transmit power to approximately $M^2/12$, while maintaining the distinguishability of x_k . In effect, all x_k 's differing by an integer multiple of M are regarded as the same symbol by the decoder as shown in figure 3. The process of encoding a symbol then becomes the process of modifying s_k to one of the equivalent representatives of the symbol. The modulo- M geometry ensures that there is one representative within distance M from every possible s_k . Tomlinson-Harashima precoder can be interpreted as a one-dimensional implementation of "writing-on-dirty-paper". This connection was first observed by Erez, Shamai and Zamir [14], who also showed that in the limit of large interference, the capacity loss when using Tomlinson-Harashima precoding is exactly the shaping loss (1.53dB). The insight is that Tomlinson precoder takes only the current s_k into account, and it does not consider future side information beyond s_k . With only causal side information, the modulo operation has to be done on a symbol-by-symbol basis, producing an output that is uniformly distributed between $-M/2$ and $M/2$. This uniform distribution corresponds to a cubic shape, thus incurring a shaping loss when compared to the spherical shape of an optimal Gaussian code. To recover the shaping loss, non-causal side information must be used to perform the modulo operation on high-dimensional spheres. As noted in [14], this is possible using a precoder based on an optimal vector quantizer.

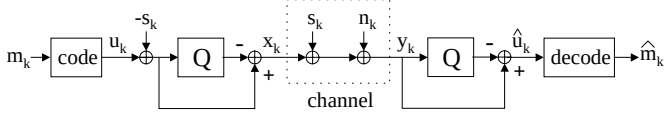


Fig. 4. Precoding via vector quantization

B. Trellis Precoding

A conceptual model for a vector quantization-based precoder is shown in figure 4. It works as follows. First, a codeword sequence is generated by an error-correcting code. The additive interference sequence $\mathbf{s}_k$ is pre-subtracted from the codeword, and the difference sequence is then quantized by a vector quantizer. The quantization noise is sent as the input to the channel. The channel adds the interference and noise. At the decoder, the received sequence is first quantized by the same quantizer. The quantization noise is sent to the decoder of the error-correcting code to recover the message. Figure 5 illustrates the relation between $\mathbf{u}_k$, $\mathbf{s}_k$ and $\mathbf{x}_k$. The spheres denote the Voronoi regions associated with the quantizer outputs. The codeword sequence $\mathbf{u}_k$ is designed to be confined within the Voronoi region. Each codeword is given multiple equivalent representatives corresponding to multiple quantizer outputs. The equivalent representatives are illustrated by the arrows in the figure. Since the interference sequence $\mathbf{s}_k$ is known non-causally, the precoder can construct an input sequence $\mathbf{x}_k$ to steer $\mathbf{s}_k$ to the closest representative of $\mathbf{u}_k$. It is easy to see that as long as the codeword is confined to the Voronoi region, perfect reconstruction is possible in the absence of noise. This vector quantization approach can be viewed as a generalization of Tomlinson-Harashima precoding. The one-dimensional modulo- M operation is replaced by a vector quantizer which does a modulo operation with respect to a Voronoi region. After the modulo operation, the precoder outputs are approximately uniformly distributed in the Voronoi region. The Voronoi region of an ideal vector quantizer is a high dimensional sphere, thus achieving a shaping gain. In other words, the transmit power may be reduced by up to 1.53dB when compared to a cubic shape.

The idea of using Voronoi region for shaping was proposed by Forney in [15], where he used a trellis code as a vector quantizer to achieve a shaping gain on an additive white Gaussian noise channel. For an AWGN channel, the capacity achieving distribution over a block length of N is a uniform distribution over an N -dimensional sphere. However, traditional rectangular constellation-based trellis codes cover the N -cube uniformly, thus suffering from a shaping loss up to 1.53dB. To recover the shaping loss, Forney proposed to first expand the constellation size slightly, then modulo the expanded sequence

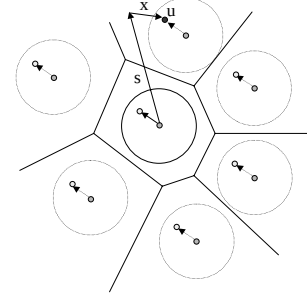


Fig. 5. Precoding with Voronoi regions

with respect to the Voronoi region of a trellis shaping code. It is clear that shaping for AWGN channel is very similar to shaping for channel with side-information. In fact, Eyuboglu and Forney further extended trellis shaping by combining shaping with one-dimensional Tomlinson precoding and trellis coding for the ISI channel [16]. This combined structure can be easily modified for channels with non-causal side information.

We now describe a trellis shaping code for channels with side information. A trellis code can be viewed as a collection of paths through the constellation. These paths have a good minimum distance property, and they can be represented by a finite state machine, thus allowing efficient decoding. Such a collection of trellis paths can also be used as reconstruction values in a vector quantizer.

The operation of the precoder is shown in figure 6. Two codes are working independently. C_c is a trellis channel code, with G_c as its k_c/n_c convolutional coset encoder. C_s is a k_s/n_s convolutional shaping code, whose Voronoi region will be used as the basis of modulo operation. The input bits are divided into two groups. First q bits are the inputs to the trellis channel code, where k_c of which are encoded by the convolutional code G_c , and the rest as uncoded bits selecting constellation points within each coset. The output of the trellis code is a constellation of size $2^{q-k_c+n_c}$. This constellation is repeated 2^{n_s} times in two dimensions resulting in 2^{n_s} non-overlapping regions. At any given time instance, one of these regions is selected by the second group of r_s inputs and the shaping convolutional code C_s . Therefore, the entire codeword can be thought of as a combination of 2^q trellis paths within each region, and the sequence of regions. Two sequences of regions are deemed equivalent if their difference is a valid codeword for C_s . The initial sequence is determined by r_s input bits and an inverse syndrome decoder H_s^{-T} for C_s . In a convolutional code, sequences differing by a valid codeword all have the same syndrome. Thus, using r_s input bits as the syndrome uniquely determines an equivalent class of n_s -bit sequences $t(D)$. Every sequence in this equivalent class can be represented by $t(D)$ plus some valid codeword in G_s . So, the shaping encoder only needs

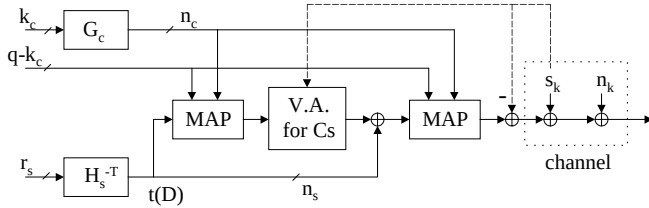


Fig. 6. Implementation of Voronoi precoder

to select the appropriate codeword to be added. The appropriate codeword is selected by the Viterbi algorithm for C_s . The criterion for such selection can in principle be anything as far as decoding is concerned. To minimize transmit power for channels with side information, the criterion here is the minimization of the square difference between the output sequence and the side information sequence. Thus, the Viterbi algorithm for C_s compares the side information sequence with the constellation sequence determined jointly by the trellis encoder C_c and the syndrome sequence for C_s , and outputs a codeword for C_s which modifies the sequence of region selects so that the resulting codeword is as close to the side information as possible. Since the side information will be added by the channel later, the encoder sends out the difference between the two to the channel. On the decoder side, the bits are recovered with a usual trellis decoder for C_c and a syndrome mapper for C_s . The decoder is identical to the one in [16].

There are two principle differences between a trellis precoder for a channel with side-information and a trellis shaping precoder for an ISI channel as described in [16]. First, a shaping code selects the region sequence so that the output codeword has the minimum energy, while a precoder selects the region sequence so that it takes the minimum energy to steer the side information to a correct codeword. The second difference is more subtle. In a shaping code for an ISI channel, a small amount of constellation expansion suffices, thus a rate 1/2 shaping code is sufficient. For precoding for a channel with side information, it is desirable to have a code rate such as 3/4 or 5/6 to force constellation expansion. The reason for expansion is to ensure that the side information sequence lies entirely within the expanded constellation. In practice where the magnitude of the side information sequence is not known in advance, it is necessary to add a modulo- M operation outside of the expanded constellation. Note that the actual transmitted constellation is not the expanded constellation, but its difference with the side information sequence. Therefore, the expanded constellation does not pose a practical concern.

The shaping gain for Voronoi precoder depends on the shape of the Voronoi region of the shaping code. Therefore, the trellis shaping codes in [15] can be used directly

for multiuser precoding with exactly the same shaping gain. In particular, a simple 4-state trellis shaping code already achieves almost 1dB shaping gain.

IV. CONCLUSION

This paper establishes a connection between the broadcast channel and the channel with non-causal transmitter side information. An achievable rate region for the non-degraded Gaussian vector broadcast channel is derived using Costa's "writing-on-dirty-paper" result for Gaussian channels with side information. This region may be approached in practice using the proposed trellis precoding scheme. The trellis precoder generalizes the Tomlinson-Harashima precoder, and it is able to pre-subtract multiuser interference completely. By utilizing non-causal side information, the trellis precoder achieves an additional shaping gain up to 1.53dB.

ACKNOWLEDGMENT

The authors wish to acknowledge Mung Chiang and Arak Sutivong for stimulating discussions.

REFERENCES

- [1] T. Cover, "Broadcast channels," *IEEE Trans. Inform. Theory*, vol. 18, no. 1, pp. 2–14, Jan. 1972.
- [2] P. Bergman, "A simple converse for broadcast channels with additive white gaussian noise," *IEEE Trans. Inform. Theory*, vol. 20, pp. 279–280, March 1974.
- [3] M. Costa, "Writing on dirty paper," *IEEE Trans. Inform. Theory*, vol. 29, no. 3, pp. 439–441, May 1983.
- [4] A.J. Goldsmith and M. Effros, "The capacity region of broadcast channels with intersymbol interference and colored Gaussian noise," *IEEE Trans. Inform. Theory*, vol. 47, no. 1, pp. 211–219, Jan. 2001.
- [5] A. Cohen and A. Lapidoth, "The Gaussian watermarking game: Part I," *submitted to IEEE Trans. Inform. Theory*, 2001.
- [6] W. Yu, A. Sutivong, D. Julian, T. Cover, and M. Chiang, "Writing on colored paper," in *Int. Symp. Inform. Theory*, June 2001.
- [7] A. Lapidoth, private communications, Feb. 2001.
- [8] K. Marton, "A coding theorem for the discrete memoryless broadcast channel," *IEEE Trans. Inform. Theory*, vol. 25, pp. 306–311, May 1979.
- [9] G. Ginis and J. M. Cioffi, "Vector DMT: a FEXT cancellation scheme," in *Asilomar Conf.*, Nov. 2000.
- [10] G. Caire and S. Shamai, "On achievable rates in a multi-antenna broadcast downlink," in *Allerton Conf. on Communications, Control and Computing*, 2000.
- [11] W. Yu and J. Cioffi, "The sum capacity of a Gaussian vector broadcast channel," *preprint*, Sep. 2001.
- [12] M. Tomlinson, "New automatic equalizer employing modulo arithmetic," *Electr. Lett.*, vol. 7, pp. 138–139, March 1971.
- [13] M. Miyakawa and H. Harashima, "A method of code conversion for a digital communication channel with intersymbol interference," *Trans. Institute of Electronic Communications Engineering of Japan*, vol. 52-A, pp. 272–273, June 1969.
- [14] U. Erez, S. Shamai, and R. Zamir, "Capacity and lattice strategies for cancelling known interference," in *Int. Symp. Inform. Theory and Its Appl.*, Nov. 2000.
- [15] G. D. Forney Jr., "Trellis shaping," *IEEE Trans. Inform. Theory*, vol. 38(II), pp. 281–300, March 1992.
- [16] M. V. Eyuboglu and G. D. Forney Jr., "Trellis precoding: combined coding, precoding and shaping for intersymbol interference channels," *IEEE Trans. Inform. Theory*, vol. 38(II), no. 2, pp. 301–314, March 1992.