On Coding for the Broadcast Phase in the Two-Way Relay Channel

Clemens Schnurr*, Tobias J. Oechtering†, Sławomir Stańczak*
*Fraunhofer German-Sino Lab for Mobile Communications
Einsteinufer 37, 10587 Berlin, Germany.
†Technical University of Berlin,
Heinrich-Hertz-Chair for Mobile Communications,
Einsteinufer 25, 10587 Berlin, Germany

***Abstract*— In this paper, we address the problem of designing practical codes for the broadcast phase in the two-way relay channel. The paper provides guidelines for the use of well-developed single-user codes in the bidirectional broadcast channel, with the goal of achieving or closely approaching the performance bounds. In case of finite alphabet channels with the channel distortion being independent of the channel input, the performance of our coding scheme only depends on the performance of the base codes on the corresponding single-user channels. If the involved single-user codes provides the best possible performance, so also does the proposed coding scheme for the bidirectional broadcast channel.**

I. Introduction

In a two-way relay channel that is of interest in this paper, two terminals exchange their messages (two-way or bidirectional communication) using a relay node that is intended solely to help the two terminals, and therefore has no own messages to transmit. We assume that there is no direct link between the terminals and the two-way communication consists of two phases: The multiple access phase and the broadcast phase. In the first phase, both terminals send their encoded messages to the relay node whose task is to reliably decode these messages. Then, during the second phase, the relay node re-encodes and transmits the decoded data to the terminals. In this paper, we focus on the second (broadcast) phase. In doing so, we neglect potential decoding errors in the first phase, which is equivalent to the assumption that the relay node has a perfect knowledge about the messages of the both terminals. This knowledge can be used to enhance the performance of the broadcast channel [1], [2], [3], [4], [5].

Reference [1] proved the capacity region for the broadcast transmission in the two-way relay channel:

Theorem 1.1: The capacity region $\mathcal{C}_{\mathrm{BC}}$ for sending two sources W_1 and W_2 over the memoryless broadcast channel, in which terminal 1 knows W_1 and terminal 2 knows W_2 is the set of all rate pairs $(R_{\overrightarrow{\mathrm{R}2}}, R_{\overrightarrow{\mathrm{R}1}})$ satisfying

$$R_{\overrightarrow{\mathrm{R}2}} \leq I(X;Y_2|U)$$
$$R_{\overrightarrow{\mathrm{R}1}} \leq I(X;Y_1|U)$$

for some probability function $p(u)p(x|u)p(y_1,y_2|x)$, where the set of the auxiliary random variable U has cardinality bounded by $|\mathcal{U}| \leq 2$.

Here $R_{\overrightarrow{\mathrm{R}2}}$ and $R_{\overrightarrow{\mathrm{R}1}}$ are the rates for transmitting independent messages from the relay to terminal 2 and 1, respectively.

This result can be interpreted as follows: The performance for each terminal is in many cases close to the performance of a single-user channel. The possible performance loss compared to the single-user channel is a potentially suboptimal channel input distribution, which may be needed to achieve some rate tuples. It is interesting to point out that in case of scalar Gaussian channels, there is no loss compared to the marginal single-user channels, as the Gaussian input distribution maximizes jointly both the mutual information expressions.

In this paper, we address the problem of designing practical codes for this communication channel, which is referred to as the *bidirectional broadcast channel*[1]. More precisely, the paper provides guidelines for the use of well-developed single-user codes in the bidirectional broadcast channel, with the goal of achieving or closely approaching the performance bounds of Theorem 1.1. In particular, in case of finite alphabet channels with the channel distortion being independent of the channel input, the performance of our coding scheme only depends on the performance of the base codes on the corresponding single-user channels. In other words, if the involved single-user codes provides the best possible performance, so also does the proposed coding scheme for the bidirectional broadcast channel. In case of arbitrary channels, however, further performance gains may be achieved by better exploiting the distortion characteristics of the channel.

The main advantage of the proposed coding scheme is that it only involves simple operations on single-user codes designed for some types of single-user channels. Consequently, from the practical point of view, the proposed scheme is very attractive, as it merely requires minor modifications of the traditional single-user coding schemes. In particular, it can be easily generalized to the case of additional soft-information. Finally, note that the assumption of channels with finite input/output alphabets is not restrictive in view of practical system im-

This work was supported by the German Ministry for Education and Research (BMBF) under Grant 01BU0680.

[1] We use the term "bidirectional" because each receiver knows the messages intended for the other receiver. From this point on, "terminal" and "user" are used interchangeably.

plementation, since in practical systems usually finite input alphabets are used; the channel output is quantized to a finite alphabet, possibly with additional soft-information attached to every symbol, which can be used in the proposed scheme as well.

II. Coding for the bidirectional broadcast channel: Symmetric marginal channels

In this section, we propose a coding scheme for the bidirectional broadcast channel. This scheme is quite general and works for several coding techniques. After introducing some notation and basic assumptions, we present an example of a concrete realization of the scheme that is based on linear block codes. Without loss of generality, in this paper, we focus on finite code lengths. An extension to infinite code length that is relevant in view of convolutional coding is straightforward, but requires more cumbersome notation.

A. Symmetric marginal channels

In this paper, we consider a broadcast channel with finite input and output alphabets. For simplicity, if the probability transition function of a broadcast channel is $p(y_1, y_2|x)$, then channels with the conditional marginal distributions $p(y_1|x)$ and $p(y_2|x)$ are referred to as *marginal channels*. Thus, the channel from the relay node to one of the terminals is a marginal channel. It is assumed that the marginal channel distortion is independent of the channel input. Furthermore, a symmetry property defined below turns out to be crucial for our analysis.

Definition 2.1 (Symmetric marginal channel): A marginal channel is called symmetric if the following holds.

(i) The channel has an output alphabet $\mathcal{Y}_i = \{0, 1, \dots, M-1, e\}$ of cardinality $M+1$ and an input alphabet $\mathcal{X} = \mathcal{Y}_i \backslash \{e\}$ of cardinality M, where e is the erasure symbol.

(ii) For some integer L, the channel output vectors of length L are given by $\boldsymbol{y}_i = \boldsymbol{x} + \boldsymbol{z}_i \in \mathcal{Y}_i^L$, where[2] $i \in \{1, 2\}$ and $\boldsymbol{z}_i \in \mathcal{Y}_i^L$ has some distribution on $\mathcal{Y}_i^L$ independent of $\boldsymbol{x} \in \mathcal{X}^L$.

A simple example of a symmetric marginal channel is the binary symmetric channel with erasure. In this section, the marginal channels are assumed to be symmetric in the sense of the above definition. Section III illustrates potential consequences of dropping this assumption. For clarity reasons, we slightly abuse the notation by using the same symbols in $\boldsymbol{y}_i = \boldsymbol{x} + \boldsymbol{z}_i$ even if we consider inputs of length $N = aL$ for some $a \in \mathbb{Z}$.

B. Encoding and decoding

Our coding scheme is based on two base codes[3] $\mathcal{C}_1$ and $\mathcal{C}_2$ being defined over the channel input alphabet $\mathcal{X}$ with encoders E_1 and E_2 and decoders D_1 and D_2. It is assumed $\mathcal{C}_i \subseteq \mathcal{X}^N, i \in \{1, 2\}$, where N is a multiple of the channel parameter L. Adequate codes and encoders can be found by simply concatenating several codewords of some given codes to generate new codes with a code length which is, for instance, the least common multiple of the length of the both codes and the channel parameter L. Both base codes may have different coding rates R_1 and R_2 (in bits per code symbol) so that the encoders

$$E_i : \{0, 1, \dots, 2^{R_i N} - 1\} \to \mathcal{C}_i \subseteq \mathcal{X}^N$$

generate codewords $\boldsymbol{c}_i \in \mathcal{X}^N, i \in \{1, 2\}$. For each code, say code $i \in \{1, 2\}$, there is a decoder

$$D_i : \mathcal{Y}_i^N \to \{0, 1, \dots, 2^{R_i N} - 1\}$$

which is assumed to decode a received word $\boldsymbol{y}_i = \boldsymbol{c}_i + \boldsymbol{z}_i$ correctly if and only if the distortion word $\boldsymbol{z}_i$ is in the set of correctable errors $\boldsymbol{z}_i \in \mathcal{E}_i^{cor}$.

Using these two base coding schemes, we can now define the coding and thereby the code $\mathcal{C}$ for the bidirectional broadcast channel as follows:

Encoding: The encoder at the relay node is a mapping

$$\psi : \{0, 1, \dots, 2^{R_1 N} - 1\} \times \{0, 1, \dots, 2^{R_2 N} - 1\} \to \mathcal{C} \subseteq \mathcal{X}^N$$

with the encoding rule given by

$$\psi(w_1, w_2) = E_1(w_2) + E_2(w_1) \,.$$

As before, the addition is symbol-wise and defined over the Abelian additive group $\mathcal{X}$. In words, the two encoders generates $\boldsymbol{c}_1$ and $\boldsymbol{c}_2$ using the encoders on the information generated by for the other user respectively, i.e. E_1 on the information w_2 and E_2 on the information w_1. The resulting codewords are added and $\boldsymbol{x} = \boldsymbol{c}_1 + \boldsymbol{c}_2$ is transmitted via the broadcast channel to user 1 and 2, which observe $\boldsymbol{y}_1$ and $\boldsymbol{y}_2$, respectively.

Decoding: The decoders at the terminals are mappings

$$\phi_1 : \{0, 1, \dots, 2^{R_1 N} - 1\} \times \mathcal{Y}^N \to \{0, 1, \dots, 2^{R_2 N} - 1\}$$

$$\phi_2 : \{0, 1, \dots, 2^{R_2 N} - 1\} \times \mathcal{Y}^N \to \{0, 1, \dots, 2^{R_1 N} - 1\}$$

The decoding rules are as follows: On receiving $\boldsymbol{y}_1$, the first user uses its perfectly known side information w_1 to generate $\boldsymbol{c}_2 = E_2(w_1)$. Then, it calculates $\hat{\boldsymbol{c}}_1 = \boldsymbol{y}_1 - \boldsymbol{c}_2$, where "$-$" denotes the addition of the inverse element as defined in the footnote before, and decodes $\hat{\boldsymbol{w}}_2 = D_1(\hat{\boldsymbol{c}}_1)$ using the decoder of the base code. The decoder of the second user is defined accordingly.

Note that both codes use the same alphabet $\mathcal{X}$. Although this assumption may appear as a significant restriction, it does not impact the generality of the analysis, as we do not require the use of all alphabet symbols. So, starting at codes defined on different alphabets with a possibly lower cardinality, the codes can be transformed to a common alphabet. The restriction is necessary and reasonable since the marginal channels use the same input.

[2]Note that $\mathcal{X}$ is an Abelian additive group $(\mathcal{X}, +)$ and the symbol-wise addition here is defined over $\mathcal{X}$. The addition is extended to $\mathcal{Y}_i, i \in \{1, 2\}$, by defining $a + e = e + a = e$ for any $a \in \mathcal{X}$. Furthermore there is no inverse element of the erasure symbol e.

[3]We use the term "code" as a set of codewords. Encoder and decoder are not part of the code, but the encoder may define the code. Encoder, decoder and code together form a coding scheme.

In fact, from practical point of view, the receiver side, i.e. the channel output alphabet, poses more problems. For instance, it is not clear what should be done if different modulations are used on a marginal channel in a real system? In this case, it may be required to restrict the alphabet at the receiver from some super-alphabet to the needed modulation alphabet before detection. A crucial point here is that the subtraction now need to be done in signal space without knowing the received constellation point. However, using a scheme similar to the proposed one, where addition is defined in an modulo like manner based on some lattice on the signal space as it is done in nonlinear pre-coding [6], [7], one arrives at a practical scheme for the broadcasting in the two-way relay channel. This is nothing but a straightforward generalization of the proposed scheme to coding in signal space, where the modulo addition is used to fulfill a possibly given power restriction. However, the Theorems of this paper concerning the performance of the codes do not generalize to such channels. The resulting coding is in general suboptimal. A more elaborated discussion on this issue is given in Section IV of this paper.

C. A conservation law for the coding scheme

The following Theorem shows that the performance characteristics of the base codes in the corresponding marginal single-user channels are preserved under the proposed coding scheme.

Theorem 2.2: The proposed coding scheme with encoder-decoder pair $(\psi, \phi_i), i \in \{1, 2\}$ and code $\mathcal{C}$ has exactly the same performance as the base coding scheme consisting of the encoder-decoder pair (E_i, D_i) and base code $\mathcal{C}_i$ in the corresponding marginal single-user channel.

Proof: User 1 observes $\boldsymbol{y}_1 = \boldsymbol{x} + \boldsymbol{z}_1$ and computes

$$\boldsymbol{y}_1 - \boldsymbol{c}_2 = \boldsymbol{x} + \boldsymbol{z}_1 - \boldsymbol{c}_2 = \boldsymbol{c}_1 + \boldsymbol{c}_2 + \boldsymbol{z}_1 - \boldsymbol{c}_2 = \boldsymbol{c}_1 + \boldsymbol{z}_1 .$$

By assumption, this can be decoded correctly iff $\boldsymbol{z}_1 \in \mathcal{E}_1^{cor}$. Now, since the distribution of $\boldsymbol{z}_1$ is the same as in the marginal channel, the performance of the code is the same as in the marginal channel. The same reasoning holds for the second user; this completes the proof. ■

Remark 2.3: The coding at the relay node can be seen as a joint network-channel coding. Instead of using the traditional network code with a modulo addition performed on data symbols as proposed e.g. by [2], [4], [5], our scheme follows the lines of the information theoretical analysis in [1]. As a result, the modulo addition is performed on the codeword symbols. An important difference to the traditional approach is that we can easily handle the case of marginal channels of different quality, as the base codes may have different coding rates. For the mentioned reasons, in what follows, the proposed scheme is referred to as *joint network-channel coding* and the code $\mathcal{C}$ as *joint network-channel code*.

Note that except for the code length, the two base codes can be chosen independently. Indeed, as far as the coding performance is concerned, it does not even matter whether they are the same or have completely different codewords. Also note that in general, the resulting encoding mapping ψ is not necessarily bijective so that there may be no way to decode messages without side information.

D. Two simple examples

As a first example, consider two identical linear codes $\mathcal{C}_1 = \mathcal{C}_2 = \hat{\mathcal{C}}$. In this case, the resulting joint network-channel code is the same as the base code so that $\mathcal{C} = \hat{\mathcal{C}}$. Only encoder ψ and the decoders ϕ_i are different as they depend on the side information. Note that without side information, no information symbol can be recovered. Due to the linearity of all the operations, the addition and subtraction can be performed on the information symbols without any performance loss. Therefore, the modulo addition on the data symbols as e.g. proposed by [2], [4], [5] can be seen as a special case of the proposed coding scheme. Another interpretation is that in this case, the side information is used for data compression. In fact, in this case, the encoding is a classical Slepian-Wolf encoding [8], where the transmitter needs to compress (w_1, w_2) for two users that have side information w_1 and w_2, respectively [9]. The compressed data is broadcasted using an error correction code, that fits for both the marginal channels. Following this interpretation there is no side information used for channel coding.

Now consider two base codes having different codewords. In this case, the resulting joint network-channel code may have more codewords than either of the two base codes. This may be for instance true if the two channels are of different quality. To illustrate this, consider the following binary codes $\mathcal{C}_1 = \{000, 111\}$ and $\mathcal{C}_2 = \{000, 110, 101, 011\}$, in which case the encoding mapping ψ is bijective, meaning that there is no compression using side information. Because of the increased number of codewords, both users may not be able to correct the transmission error without side-information; for the example at hand, all tuples in $\{0, 1\}^3$ are codewords, and therefore the error correction or detection is impossible without side information. Error correction decoding (or error detection for the second user) becomes possible by restricting to a subset of codewords with the help of the side information. In this example, there is no compression at all since $|\mathcal{C}| = |\mathcal{C}_1||\mathcal{C}_2|$, but the side information is used in a similar way as the already decoded information is used in the decoding of the weaker signals in interference cancellation schemes, e.g [10].

E. Interpretation and remarks

In general, the joint network-channel code utilizes more codewords than the users are able to differentiate. Furthermore, one has $\max\{|\mathcal{C}_1|, |\mathcal{C}_2|\} \leq |\mathcal{C}| \leq |\mathcal{C}_1||\mathcal{C}_2|$, where strict inequalities hold for many base codes, and therefore the proposed scheme is in general neither full compression nor pure channel coding. In fact, depending on the choice of the base codes, one achieves some point on a tradeoff curve between these two extreme points. By eliminating all the effects caused by the simultaneous transmission, the proposed joint network-channel coding restricts the error correction decoding at the receiver to a sub-code, which is the corresponding base code. By the symmetry of the channel, the decoder of this error correction code "sees" only the distortion caused by the marginal channel. As the code is designed for this channel, the performance will be as desired.

Another interpretation of the coding/decoding scheme, which is more along the lines of the information theoretical result is that of a set of codes[4]. The direct part of the proof of Theorem 1.1 in [1] indicates that one should design several codes with encoder and decoder for each user, all corresponding to the marginal channel; one triplet of encoder, code and decoder for each potential side information.

Following these requirements, we need $2^{R_1 N}$ encoders E_{1,w_1} to encode w_2 and $2^{R_2 N}$ encoders E_{2,w_2} to encode w_1. Furthermore, only one codeword should be transmitted. Therefore the encoders and codes need to be interwoven with each other such that $E_{1,w_1}(w_2) = E_{2,w_2}(w_1)$. If this equation holds then given (w_1, w_2) the output of both encoders E_{1,w_1} and E_{2,w_2} is the same, even though the codeword carries different information for different receivers.

In general, the requirements on coding that follow from Theorem 1.1 seem to be hard to satisfy. Nevertheless, the proposed coding scheme is a simple recipe to realize codes with the required properties: The codes and corresponding encoder mappings for the different side information are created by shifting one base code and the encoder mapping in $\mathcal{X}^N$. The shift is realized by adding the encoded side information. The resulting codes are coset-codes of the original code. By the commutativity of the addition, this design guarantees the desired interconnection $E_{1,w_1}(w_2) = E_{2,w_2}(w_1)$ of the codes. At the decoder, instead of decoding the shifted code, we can as well invert the shift and decode the base codes. If needed, this also indicates how soft-information should be handled: Just shift the soft-information in the same way as the code.

F. An optimality property of the scheme

An immediate consequence of Theorem 2.2 is that in our setting, the proposed coding scheme performs as close to the optimal performance as the underlying codes do for the marginal channels. The following Theorem proves an optimality property of the proposed procedure in the sense that it enables us to construct a coding scheme whose performance for both users is at least as good as any other coding scheme for the bidirectional broadcast channel.

Theorem 2.4: For the considered symmetric channel, we can always find two base codes of length N with encoders $E_1(w_2)$ and $E_2(w_1)$ and decoders $D_1(\boldsymbol{y}_1), D_2(\boldsymbol{y}_2)$ such that the resulting code with encoder $\psi(w_1, w_2)$ and decoders $\phi_1(w_1), \phi_2(w_2)$ has a probability of error for both the users at least as low as any given code of length N with encoder $\theta(w_1, w_2)$ as well as decoders $\rho_1(w_1, \boldsymbol{y}_1)$ and $\rho_2(w_2, \boldsymbol{y}_2)$. This is still possible, even if a given code with encoder $\theta(w_1, w_2, q)$ permits a certain randomness depending on some randomization parameter $q \in \mathcal{Q}$ as long as q is independent of w_i.

Proof: Suppose that $q \in \mathcal{Q}$ is a randomization parameter independent of w_1 and w_2 and that some code for a symmetric channel with encoder $\theta(w_1, w_2, q)$ is given. For deterministic encoding, let the cardinality of $\mathcal{Q}$ be $|\mathcal{Q}| = 1$. Based on this code and encoder, we can define codes for the marginal channels of both the users by fixing the side information and the randomization parameter, until we find good marginal codes. More precisely: For an arbitrary pair (w_1, q) of side information w_1 and the randomization parameter q, define a code for user 1 with encoder $\hat{E}_{1,w_1,q}(w_2) = \theta(w_1, w_2, q)$ and decoder $\hat{D}_{1,w_1,q}(\boldsymbol{y}_1) = \rho_1(w_1, \boldsymbol{y}_1, q)$. This encoder defines a code of length N. Check the probability of error for this coding scheme on the marginal channel. If the probability of error of this code and this user is at least as low as that of the given code, fix encoder $E_1(w_2) = \hat{E}_{1,w_1,q}(w_2)$ and decoder $D_1(\boldsymbol{y}_1) = \hat{D}_{1,w_1,q}(\boldsymbol{y}_1)$, and then use this code as base code for user 1. Otherwise, check another pair (w_1, q) of side information w_1 and randomization parameter q. As q is independent of w_1 and w_2, and the decoder is the same as in the original code, there exists at least one such code. The same should be repeated for user 2. By Theorem 2.2, the code with encoder $\psi(w_1, w_2)$ constructed based on $E_1(w_2)$ and $E_1(w_2)$ has an error probability at least as low as the given code for both the users, if the decoding $\phi_i(w_i)$ use $D_i(\boldsymbol{y}_i)$ after subtracting the encoded side information. Furthermore, the code is of length N as required by the Theorem. ■

This proves the optimality of the proposed scheme. By using the coding not on the data but, as proposed by the information theoretical result, on the channel input, we can use different coding rates for both the users and may — provided that we find good single user codes — achieve any point in the capacity region given by Theorem 1.1 proved in [1].

G. Example

As an example for the proposed coding scheme, we present a code construction starting from two linear block-codes of length N. In this case, the encoders E_1 and E_2 of the codes are given as two generator matrices $\boldsymbol{G}_1$ and $\boldsymbol{G}_2$ and the codewords are calculated e.g. for the first user by $\boldsymbol{c}_1 = \boldsymbol{w}_2 \boldsymbol{G}_1$ from data words $\boldsymbol{w}_i \in \mathcal{X}^{R_i N}$, where the coding rates R_i are given in information symbols per code symbol.

Assume a binary Hamming code for the first user with the generator matrix given by

$$\boldsymbol{G}_1 = \begin{bmatrix} 1001110 \\ 0101011 \\ 0011101 \end{bmatrix}$$

which has a minimum code distance of 3 and a code for user 2 with the generator matrix given by

$$\boldsymbol{G}_2 = \begin{bmatrix} 1010101 \\ 0101011 \end{bmatrix}$$

which has a minimum distance of 4. Furthermore, note that the corresponding check matrices $\boldsymbol{H}_i$ satisfy $\boldsymbol{G}_i \boldsymbol{H}_i^T = 0$ so that they are the generator matrices of the dual codes.

[4]Note that the codes in the set need not to be disjoint. It may even happen that some of these codes are identical and only the decoder and encoder differ. In fact, the design is a classical coset code design as proposed by [11], except that our scheme may use the same coset with different mappings several times.

In the case of linear block codes, the encoding at the relay is simply

$$\boldsymbol{x} = [\boldsymbol{w}_2 \boldsymbol{w}_1] \begin{bmatrix} \boldsymbol{G}_1 \\ \boldsymbol{G}_2 \end{bmatrix}$$

We do not require the rows of $\boldsymbol{G}_1$ and $\boldsymbol{G}_2$ to be linearly independent. In the example, the overall code has a minimum distance of 2. For instance, the codeword 1010101 corresponds to $w_2 = 000$ and $w_1 = 10$, while 1010011 is the codeword for $w_2 = 101$ and $w_1 = 00$. However, due to subtracting the codeword generated from the side information, the codes used for decoding at the terminal have a code distance of 3 and 4, respectively. This is because all the codeword pairs with distance 2, as those in the example above, belong to different cosets for both users, which again is indicated by the different side information.

Error detection can be performed via syndrome calculation. If this is done before subtracting the encoded side information, this reveals the coset structure on which the scheme is based: The first terminal uses the side information to calculate the new syndrome for error-free transmission $\boldsymbol{s}_1 = \boldsymbol{w}_1 \boldsymbol{G}_2 \boldsymbol{H}_1^T$. Now use the matrix $\boldsymbol{H}_1$ to check the received codeword $\boldsymbol{y}_1$ before subtraction with this syndrome $\boldsymbol{y}_1 \boldsymbol{H}_1^T \stackrel{?}{=} \boldsymbol{s}_1$. If the equation is fulfilled, there is no error; if not, the difference between the left and right side may be used to locate the error. Error detection at terminal 2 works accordingly. As an alternative to subtracting the encoded side information, the decoding can shift the mapping defined by the generator matrix $\boldsymbol{G}_i$ according to the coset indicated by the side information. Note that now not only the syndrome but all the information need to be taken into account: The coset-code may have an all-zero-syndrome, as for example the coset of code $\mathcal{C}_{1,\boldsymbol{w}_1}$ for user 1 with side information $\boldsymbol{w}_1 = 01$. But the mapping that is used in the encoding is different from the original mapping given by $\boldsymbol{G}_1$.

III. Non-symmetric channels

In the last section, the case of symmetric channels was considered. This does not include a channel in which the channel distortion word $\boldsymbol{z}_i$ depends on the channel input. A simple counter example presented below shows that in this case, the performance of the base codes may degrade if we use the proposed scheme.

Assume a binary symmetric channel to user 2 with crossover probability such that a three bit repetition code achieves the desired probability of error, i.e. we have $\mathcal{C}_2 = \{000, 111\}$. The non-symmetric channel to user 1 outputs 3 random bits, whenever the binary sum of the input of three bits equals 1. If the sum is 0, the channel is error-free. A 3-bit code for this channel may simply use all the four even weight sequences of length 3 and will achieve zero error probability. Therefore, a possible code is $\mathcal{C}_2 = \{000, 011, 110, 101\}$. The joint network-channel code resulting from the proposed scheme is now $\mathcal{C} = \{0,1\}^3$. If we use this code to broadcast the information to both users, user 1 cannot achieve the desired performance. Half of the codewords of this code are inappropriate for the channel to user 1. Whenever $\boldsymbol{c}_2 = 111$, the output of the channel to user 1 is random. Even with side information, user 1 cannot decide which data word is correct. This is because the coset code $\mathcal{C}_{1,w_1} = \{111, 100, 001, 010\}$ for this side information does not perform as desired for the given channel. The second user will still achieve the desired performance as its channel is symmetric.

The coding may be changed such that the performance of user 1 is increased at the cost of user 2. As the base code $\mathcal{C}_1$ of user 1 is the only set of codewords of cardinality 4 that has the desired performance, in order to protect the performance of user 1, the new base code $\tilde{\mathcal{C}}_2$ for user 2 needs to consist only of codewords of the base-code $\mathcal{C}_1$. This is sufficient to obtain coset codes such that $\mathcal{C}_{1,w_1} = \mathcal{C}_1$ for any w_1 as the code $\mathcal{C}_1$ is a linear code. Therefore, $\tilde{\mathcal{C}}_2$ may consist of any two codewords of even weight and three bit length, as e.g. $\tilde{\mathcal{C}}_2 = \{000, 011\}$. This degrades the performance of user 2 as the code distance now is at most 2. User 2 may as well use up to 4 codewords without further degradation of its performance.

A similar tradeoff effect can be noticed in the information theoretical result of Theorem 1.1: Whenever the marginal channels do not match concerning their optimizing input distribution, there is a tradeoff between the performance of the two users. In the proposed coding scheme the needed match is that of cosets of codes: The coset codes, which are created by the side information, need to match the channel such that, on average, the desired performance is achieved. By the symmetry assumption, we assure that all cosets have the same performance. In case of non-symmetric channels, an optimal code design might become far more difficult. A simple but potentially suboptimal solution is to combine the coding with a pseudo noise sequence known to transmitter and receiver. By randomly scrambling the channel input, the channel seen by the code becomes symmetric. In doing so, one looses the possibility to use certain structure of the channel noise to enhance the decoding performance.

IV. Additiv noise channels with non-discrete alphabets

In practice, wireless channels are not constrained to have finite discrete input and output alphabets. In fact, their inputs and outputs can have arbitrary values and channel distortion often can be modeled in terms of additive noise. In this section, we briefly consider such channels.

First of all note that as far as the results of this paper are concerned, we do not need the restriction to discrete finite alphabets. In fact, the following two requirements must hold: 1. There must be an Abelian additive group $(\mathcal{X}, +)$ defined on the input and the output alphabet. This is often fulffilled e.g. by channels that have input and output in $\mathbb{C}$ or $\mathbb{R}$. 2. the channel distorion is additive noise only, which is independent of the channel input and is added according to the definition of $(\mathcal{X}, +)$. These requirements hold for AWGN channels as well as for many more general additive noise channels, where the noise is not neccessarily Gaussian.

For a channel that fulfills these two conditions, all the results of this paper hold, provided that the described coding is performed in signalspace, that is, on the channel input/output alphabet according to $(\mathcal{X}, +)$. The problem is that, in practice, we have some additional constraints such as power constraints on the channel input. The use of the coding scheme is in general suboptimal if e.g. power constraints are given. In fact, compared to the point-to-point transmission on the marginal channels for which the base codes are designed and that serve as performance-reference for the results in this paper, the required transmit power is increased by the use of the coding scheme. Therefore the scheme is not optimal in view of the information theoretical result of Theorem 1.1 adapted to this case by restricting the probabiliy disitribution $p(x)$ to fulfill the power constraint. The information theoretical result states, that we should be able to achieve for both users the performance of some[5] transmission on the marginal channels using the same power as in the broadcast channel.

Now the question is whether there is a simple solution to the coding problem in additive noise channels? From a practical point of view, a good but in general suboptimal solution could be as follows: Define a new Abelian additive group $(\hat{\mathcal{X}}^T, \oplus)$ where $\hat{\mathcal{X}}^T$, $T \in \mathbb{Z}$ is such that every $\hat{\boldsymbol{x}}^T \in \hat{\mathcal{X}}^T$ and every sequence of $\hat{\boldsymbol{x}}^T$ fulfills the given constraint. For example given a power constraint $E\{|x|^2\} \leq P$ and $\mathcal{X} \subset \mathbb{C}$, one may use a modulo like operation on real and imaginary part such that $\Re(x) \in [-\sqrt{\frac{P}{2}}, +\sqrt{\frac{P}{2}}]$ and $\Im(x) \in [-\sqrt{\frac{P}{2}}, +\sqrt{\frac{P}{2}}]$; now every symbol $x \in \hat{\mathcal{X}}$ fulfills the constraint[6]. More advanced schemes may consider more than one symbol $T > 1$ to fulfill the constraints or use a more complex group. In addition, one needs two base codes with $\mathcal{C}_i \in \hat{\mathcal{X}}^{nT}$ for some $n \in \mathbb{Z}$. Now add the two codewords $\boldsymbol{c}_1$ and $\boldsymbol{c}_2$ according to the Abelian additive groupe $(\hat{\mathcal{X}}^T, \oplus)$. The resulting symbol sequence fulfills the given constraints.

The next problem to cope with is that the noise is not added in accordance with $(\hat{\mathcal{X}}^T, \oplus)$. Furthermore, noise e.g. Gaussian noise is not restriced to $\hat{\mathcal{X}}$, and therfore decoding must be performed in the unconstraint receive alphabet $\mathcal{Y}$. As in the discussion of the last sections, the decoder now may decode the coset code induced by the side information. None the less, inverting the shift from the base code to the coset code is in general not a simple substraction of the encoded data, as the noise is not added according to $(\hat{\mathcal{X}}^T, \oplus)$. Therefore decoding may be more complex than before. In particular, the base decoder may be not applicable. Furthermore, not all coset codes may achieve the same performance as the base code. Therfore the results of this paper do not apply for this class of channels, if a constraint to the input of the channel is given.

As a toy example, consider an AWGN channel with real input and real output and a constraint on the amplitude per symbol $x \in [-A, A)$. Following the above proposal, we use $\mathcal{C}_1 = \{-A, +A - \epsilon\}, \epsilon > 0$ and $\mathcal{C}_2 = \{-A, 0\}$ and add the encoded symbols according to $c = ((c_1 + c_2 + A) \mod 2A) - A$. The code of user 1 has a good performance for $c_2 = 0$, but for $c_2 = -A$ we get the coset code $\mathcal{C}_{1,c_2=-A} = \{0, -\epsilon\}$, which has a bad performance if ϵ is small.

Obviously, this example is constructed to fail; for application in real systems, however, the proposed coding in signal space may often be a good option. For the optimal performance, joint tranmission schemes need to be developed.

V. Conclusion

In this paper, we proposed a framework for designing coding schemes for the symmetric bidirectional broadcast channel. Our framework is optimal in the sense that for every given coding scheme, we can construct a new coding scheme, which performs at least as good as the given one. Furthermore, the scheme has the advantage that we can use well-developed single user codes; there is no need to develop special codes for the symmetric bidirectional broadcast channel.

[5] Note that in the general case, the performance of the set of codes used on the marginal channel may be suboptimal for that single user channel.

[6] There may be symbols $x \notin \hat{\mathcal{X}}$ that fulfill the constraint as well.

References

[1] T. J. Oechtering, C. Schnurr, I. Bjelakovic, and H. Boche, "Achiebable rate region of a tow phase bidirectional relay channel," in *Proceedings of the 41st Annual Conference on Information Sciences and Systems (CISS)*, Mar 2007, accepted.

[2] Y. Wu, P. A. Chou, and S. Y. Kung, "Information exchange in wireless networks with network coding and physical-layer broadcast," in *Proceedings of the 39th Annual Conference on Information Sciences and Systems (CISS)*, March 2005.

[3] B. Rankov and A. Wittneben, "Spectral efficient signaling for half-duplex relay channels," in *Asilomar Conference on Signals, Systems, and Computers 2005*, Nov. 2005.

[4] R. W. Yeung, "Network Coding Theory Part I/II: Single/Multiple Source," *Foundations and Trends in Communications and Information Theory*, vol. 2, no. 4–5, pp. 241–381, 2005.

[5] C. Fragouli, J. Y. Boudec, and J. Widmer, "Network coding: An instant primer," *ACM SIGCOMM Computer Communication Review*, vol. 36, no. 1, pp. 63–68, Jan. 2006.

[6] M. Tomlinson, "New automatic equaliser employing modulo arithmetic," *Electronics Letters*, vol. 7, no. 5,6, pp. 138–139, Mar. 1971.

[7] H. Harashima and H. Miyakawa, "Matched transmission technique for channels with intersymbol interference," *IEEE Trans. Commun.*, vol. COM-20, pp. 774–80, Aug. 1972.

[8] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inform. Theory*, vol. IT-19, no. 4, pp. 471–480, Jul 1973.

[9] J. K. Wolf, "Source coding for a noiseless broadcast channel," in *Proceedings of the 38th Annual Conference on Information Sciences and Systems (CISS)*, 2004, pp. 666–671.

[10] G. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multiple antennas," *Bell Laboratories Technical Journal*, vol. 1, no. 2, pp. 41–59, 1996.

[11] S. S. Pradhan and K. Ramchandran, "Distributed source coding using syndromes (DISCUS): Design and construction," *IEEE Trans. Inform. Theory*, vol. 49, no. 3, pp. 626–642, Mar 2003.