Carbon Copying Onto Dirty Paper

Ashish Khisti and Gregory Wornell
MIT
Dept. EECS
Cambridge, MA, 02139, USA
Email: {khisti,gww}@mit.edu

Uri Erez
Tel Aviv University,
Dept. of Electrical Enineering
Ramat Aviv, 69978, Israel
Email: uri@eng.tau.ic.il

Amos Lapidoth
ETH Zürich
ISI (D-ITET), ETH Zentrum
CH-8092 Zürich, Switzerland
Email: lapidoth@isi.ee.ethz.ch

***Abstract*— A generalization of the problem of writing on dirty paper is considered in which one transmitter sends a common message to multiple receivers. Each receiver experiences on its link an additive interference, which is known noncausally to the transmitter but not to any of the receivers.**

In this work, we focus on the Gaussian case with two users and independent interferences. We provide upper and lower bounds on capacity. At high interference-to-noise ratios, we show that time-sharing is (asymptotically) optimal. This settles the conjecture by Steinberg and Shamai [10]. At high signal-to-noise ratios, we propose a superposition dirty paper code that achieves within $1/4$ bit/symbol of capacity. An extension to the case of correlated interferences is also discussed.

I. Introduction

The study of communication over channels controlled by a random state parameter known only to the transmitter has received renewed attention due to emerging applications such as digital watermarking and broadcasting over multi-input-multi-output (MIMO) channels. The case of point to point channels with random parameters has been studied in [1], [3]–[5], [9] in several different scenarios. In particular, Costa [1] considers a model in which there is an additive white Gaussian interference ("dirt"), which constitutes the state, in addition to independent additive white Gaussian noise. The key result in this "dirty paper coding" scenario is that there is no loss in capacity if the interference is known only to the transmitter.

This paper examines the *common-message* broadcast channel, which we refer to as the *multicast* channel. Specifically, we consider a scenario in which one transmitter broadcasts a common message to multiple receivers over Gaussian channels. In the case of interest, associated with the link to each receiver is a corresponding additive white Gaussian interference, in addition white Gaussian noise. The transmitter has perfect noncausal knowledge of all these interference sequences, but none of the receivers have knowledge of any of them. This model and its generalizations arise rather naturally not only in a variety of multi-antenna wireless multicasting problems, but also in applications of dirty paper coding where only imperfect knowledge of the state is available to the transmitter.

The capacity of some binary versions of such multicast channels is reported in [6], [8]. For more general channels, [10] reports achievable rates for broadcasting common and independent messages over a discrete memoryless broadcast channel with noncausal state knowledge at the transmitter. The case of two-user Gaussian channels with jointly and individually independent identically distributed (i.i.d.) Gaussian interferences on each link is also considered in [10], for which it is conjectured that in the limit of large interference, time-sharing between the two receivers is optimum even when both are only interested in a common message. In this paper we establish that this conjecture is true. We upper bound the capacity of the Gaussian channel and show that it approaches the time-sharing rate in this limit. In addition, we also present a coding scheme that achieves within 1/4 bit/symbol of the capacity[1] in the high signal to noise ratio (SNR) limit for any interference power.

This work was supported in part by NSF Grant No. CCF-0515109, and by the HP/MIT Alliance.

II. Channel Model

The two user memoryless Gaussian case is depicted in Fig. 1. It consists of one transmitter and two receivers with an additive white Gaussian noise of variance N on each link. In addition, each link also experiences an additive white Gaussian interference S_k^n, $k = 1, 2$ of variance Q. Unless otherwise stated, we assume that the two interference sequences are also mutually independent. Thus, the observation at receiver k takes the form

$$Y_k^n = X^n + S_k^n + Z_k^n, \qquad k = 1, 2. \tag{1}$$

The transmitted sequence X^n is a function of the message W and the state sequences S_1^n,S_2^n. It satisfies the power constraint

$$E\left[\frac{1}{n}\sum_{i=1}^{n} X_i^2(W, S_1^n, S_2^n)\right] \leq P, \tag{2}$$

where the expectation is over the message and the state sequences. Finally, note that without loss of generality, we may set $N = 1$, and interpret P as the signal-to-noise ratio (SNR), and Q as the interference-to-noise ratio (INR).

III. Capacity Bounds

For this channel, we have the following bounds on capacity.

Theorem 1: The Gaussian multicast channel capacity is bounded according to:

$$R_- \leq C \leq R_+, \tag{3}$$

[1]Throughout this work, symbol refers to a *real* symbol.

1-4244-0092-9/06/$20.00

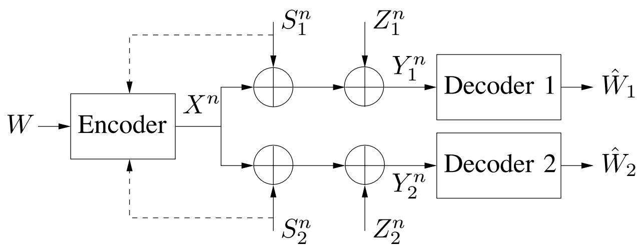

Fig. 1. Two-user Gaussian multicast channel model with additive interference. The encoder maps message W into codeword X^n. The state takes the form of interference sequences S_1^n and S_2^n. Each channel output $Y_k^n = X^n + S_k^n + Z_k^n$ is decoded to produce message estimate $\hat{W}_k$.

where[2]

$$R_+ = \begin{cases} \frac{1}{4}\log(1+P) + \frac{1}{4}\log\left(\frac{P+Q+1+2\sqrt{PQ}}{Q}\right) & Q \geq 4 \\ \frac{1}{4}\log\left(\frac{1+P}{Q/4+1}\right) + \frac{1}{4}\log\left(\frac{P+Q+1+2\sqrt{PQ}}{Q/4+1}\right) & Q < 4 \end{cases} \tag{4}$$

and

$$R_- = \begin{cases} \frac{1}{2}\log\left(1+\frac{P}{Q/2+1}\right) & Q/2 < 1 \\ \frac{1}{2}\log\left(\frac{P+Q/2+1}{Q}\right) + \frac{1}{4}\log\left(\frac{Q}{2}\right) & 1 \leq Q/2 < P+1 \\ \frac{1}{4}\log(1+P) & Q/2 \geq P+1. \end{cases} \tag{5}$$

Proof: A proof for the upper bound is presented in the full paper [7]. Our approach involves bounding the rate of the two-interference channel in terms of that for a single-interference channel. The lower bound (5) is an explicit expression of the following maximization:

$$R_- = \max_{\{(P_1,P_2): P_1 \geq 0, P_2 \geq 0, P_1+P_1 \leq P\}} R(P_1, P_2) \tag{6a}$$

with

$$R(P_1,P_2) \triangleq \frac{1}{2}\log\left(1+\frac{P_1}{P_2+Q/2+1}\right) + \frac{1}{4}\log\left(1+P_2\right). \tag{6b}$$

Accordingly, we show the achievability of (6b). The proposed scheme, combines superposition coding, dirty paper coding, and time-sharing, and exploits a representation of the interferences in the form

$$\begin{aligned} S_1^n &= S^n + V^n \\ S_2^n &= S^n - V^n, \end{aligned} \tag{7}$$

where

$$\begin{aligned} S^n &= (S_1^n + S_2^n)/2 \\ V^n &= (S_1^n - S_2^n)/2. \end{aligned} \tag{8}$$

Specifically, the encoding is as follows:

[2] We can also tighten the upper bound further by considering $\min\left(R_+, \frac{1}{2}\log\left(1+P\right)\right)$, where the second expression corresponds to the multicasting rate when the interference is absent.

1) Decompose the message W into two submessages W_S and W_V and divide the power P into two powers P_1 and P_2 so that $P = P_1 + P_2$.
2) Use dirty paper coding [1] to (independently) encode message W_S into a codeword U_S^n in a (random i.i.d. Gaussian) codebook $\mathcal{C}_S$ using power P_1 for a channel with interference S^n and Gaussian noise of power $P_2 + Q/2 + 1$. Transmit $X_S^n = U_S^n - \alpha_S S^n$, where $\alpha_S = P_1/(P+Q/2+1)$. The corresponding rate is
$$R_S = \frac{1}{2}\log\left(1+\frac{P_1}{P_2+Q/2+1}\right). \tag{9}$$
3) Use dirty paper coding to encode message W_V into a codeword U_V^n in a codebook $\mathcal{C}_V$ using power P_2 for a channel with interference either $(1-\alpha_S)S^n + V^n$ or $(1-\alpha_S)S^n - V^n$, and Gaussian noise of power 1. Transmit the corresponding $X_V^n = U_V^n - \alpha_V\{(1-\alpha_S)S^n + V^n\}$ or $X_V^n = U_V^n - \alpha_V\{(1-\alpha_S)S^n - V^n\}$, with $\alpha_V = P_2/(P_2+1)$. Time-share evenly between these two possibilities. In each case the corresponding rate is
$$R_V = \frac{1}{2}\log(1+P_2). \tag{10}$$
4) Send the superposition $X^n = X_S^n + X_V^n$, which has power P over the channel.

The decoding exploits successive cancellation (stripping) and proceeds as follows:

1) Decode U_S^n from Y_1^n or Y_2^n treating X_V^n as part of the noise. The received signals are of the form
$$\begin{aligned} Y_1^n &= X_S^n + S^n + (V^n + Z_1^n + X_V^n) \\ &= U_S^n + (1-\alpha_S)S^n + (V^n + Z_1^n + X_V^n), \\ Y_2^n &= X_S^n + S^n + (-V^n + Z_2^n + X_V^n) \\ &= U_S^n + (1-\alpha_S)S^n + (-V^n + Z_2^n + X_V^n). \end{aligned}$$
The rate R_S in (9) ensures that the resulting $\hat{W}_S$ equals W_S with high probability.
2) Subtract the decoded U_S^n from each of Y_1^n and Y_2^n, so that the residual signals $\tilde{Y}_i^n = Y_i^n - U_S^n$ are of the form
$$\tilde{Y}_1^n = X_V^n + ((1-\alpha_S)S^n + V^n) + Z_1^n, \tag{11}$$
$$\tilde{Y}_2^n = X_V^n + ((1-\alpha_S)S^n - V^n) + Z_2^n. \tag{12}$$
The rate R_V in (10) ensures that X_V^n can be decoded from either $\tilde{Y}_1^n$ or $\tilde{Y}_2^n$ so that the resulting $\hat{W}_V$ equals W_V with high probability at the corresponding receiver. Specifically, for the fraction of time that the transmitter encodes W_V for interference $(1-\alpha_S)S^n + V^n$, user 1 can recover W_V, while for the fraction of time that the transmitter encodes W_V for interference $(1-\alpha_S)S^n - V^n$, user 2 can recover W_V.

From this coding strategy, we see that the average rate delivered to each receiver is identical, i.e., $R_S + (1/2)R_V$. Maximizing this rate over the choices of P_1 and P_2 subject to the constraint $P = P_1 + P_2$ optimizes the lower bound, whence (6a).

■

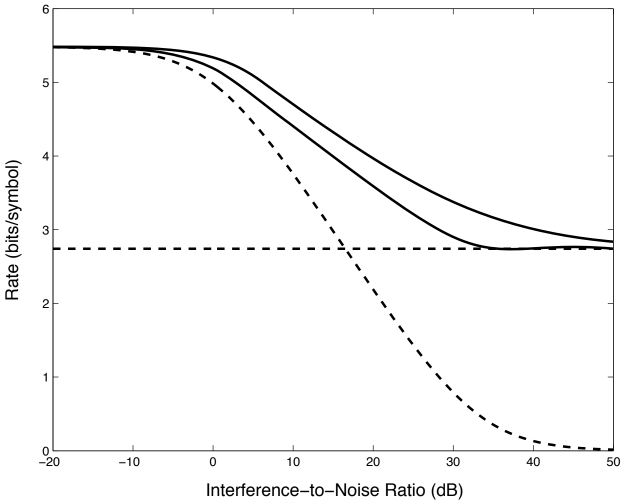

Fig. 2. Upper and lower bounds on the capacity of the two-user Gaussian multicast channel, as a function of INR Q for an SNR $P = 33$ dB. The solid curves depict the two bounds of (3). The horizontal dashed line indicates the performance of time-sharing, while the other dashed curve indicates the performance of a strategy in which the side information is treated by the transmitter as additional noise on each link.

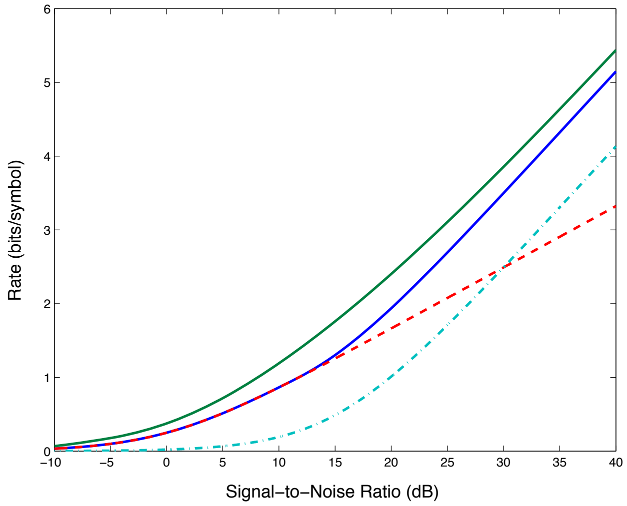

Fig. 3. Upper and lower bounds on the capacity of the two-user Gaussian multicast channel, as a function of SNR P for an INR $Q = 15$ dB. The solid curves depict the two bounds of (3). The dashed curve indicates the performance of time-sharing, while the dash-dotted curve indicates the performance of a strategy in which the side information is treated by the transmitter as additional noise on each link.

IV. Discussion

From (5), we obtain several useful insights. First, note that when $Q/2 \geq P+1$, our lower bound reduces to time-sharing, while if $Q/2 \leq 1$ it reduces to dirty paper coding with respect to S^n. In the moderate interference regime, our bound shows that one can generally achieve a gain over these two strategies by a superposition coding approach that combines them.

The behavior of the bounds as a function of INR is depicted in Fig. 2 for a fixed SNR of $P = 33$ dB. When the INR is very small, Fig. 2 reflects that the side information can be ignored by the transmitter without sacrificing rate. Similarly, when the INR is large, Fig. 2 reflects that time-sharing is capacity-achieving. Both these observations can be quantified.

For the low INR behavior, it suffices to note from (4) that

$$\lim_{Q\to 0} C \leq \lim_{Q\to 0} R_+ = \frac{1}{2}\log(1+P), \tag{13}$$

which is the same as $\lim_{Q\to 0} R_{\mathrm{IS}}$, where R_{IS} is the rate achieved by ignoring the side-information, i.e.,

$$R_{\mathrm{IS}} = \frac{1}{2}\log\left(1+\frac{P}{Q+1}\right). \tag{14}$$

For the high INR behavior, it is suffices to note from (4) that

$$\lim_{Q\to\infty} C \leq \lim_{Q\to\infty} R_+ = \frac{1}{4}\log(1+P), \tag{15}$$

which can be achieved by time-sharing between the two users and doing Costa dirty paper coding for each user being served. We note that this result settles the conjecture made in [10].

The behavior of the bounds as a function of SNR is depicted in Fig. 3 for a fixed INR of $Q = 15$ dB. When the SNR is large, Fig. 2 reflects that having the transmitter ignore the side information does not achieve a rate particularly close to capacity, but the superposition dirty paper coding strategy corresponding to our lower bound does. More generally, the difference between the upper bound and the lower bound can be bounded as follows:

$$\lim_{P\to\infty}(C-R_-) \leq \lim_{P\to\infty}(R_+ - R_-) = \begin{cases} \frac{1}{4} & Q \geq 4 \\ \frac{1}{4}\log\left(\frac{2Q}{(1+Q/4)^2}\right) & 2 \leq Q < 4 \\ \frac{1}{2}\log\left(1+\frac{Q}{4+Q}\right) & Q < 2 \end{cases} \tag{16}$$

Hence, for INR above 3 dB, our superposition dirty paper code achieves within $1/4$ bit/symbol of capacity, while for INR below 3 dB, the superposition part of the code does not help, but the dirty paper code alone achieves within $\frac{1}{2}\log(4/3) = 0.2075$ bit/symbol of capacity. It is straightforward to verify (16); a proof can be found in the full paper [7].

We conclude this section with a few additional observations.

Some Further Remarks:

1) Feedback does not help much. An upper bound when the transmitter has perfect causal feedback from the receiver is given by

$$R_+^{\mathrm{F}} = \frac{1}{4}\log(1+P) + \frac{1}{4}\log\left(\frac{P+Q+1+2\sqrt{PQ}}{Q/2+1}\right). \tag{17}$$

Thus in the limit $Q \to \infty$, feedback can gain by no more than 1/4 bit/symbol over time-sharing.

2) Correlation among noise sequences does not matter, i.e., the upper bound in Theorem 1 is valid even when the noises Z_1^n and Z_2^n are not independent. The argument

is analogous to that for the standard broadcast channel (e.g. [2, Ch. 14]).

A. Extensions for Robust Dirty Paper Coding

Consider the a memoryless Gaussian point-to-point channel model with output

$$Y^n = X^n + S^n + Z^n, \tag{18}$$

where X^n is the channel input subject to power constraint P, S^n is a white Gaussian interference sequence of power Q not known to decoder, and Z^n is a white Gaussian noise sequence of unit power. When the interference S^n is perfectly known to the encoder, Costa's dirty paper coding is capacity achieving. However, in many applications, only imperfect knowledge of S^n is available to the encoder. One special case is the case of *causal* knowledge considered by Shannon. Another is the case of *noisy* noncausal knowledge. For these kinds of generalizations, there is interest in understanding the capacity of such channels and the structure of the associated capacity-achieving codes, which we refer to as *robust* dirty paper codes.

It is often natural to analyze such problems via their equivalent Gaussian multicast model. As an illustration, suppose that the interference in (18) is of the form $S^n = \beta S_0^n$ where $S_0^n \sim \mathcal{N}(0, Q\mathbf{I})$ is known to the encoder but β is not. Then if β is from a finite alphabet (or can be approximated as being so), i.e., $\beta \in \{\beta_1, \beta_2, \ldots, \beta_K\}$, the problem is equivalent to a Gaussian multicast problem with K users where the interference for the kth user is $\beta_k S_0^n$.

From this example it is apparent that for at least some applications, there is a need to accommodate *correlated* interferences in the Gaussian multicast model. While a thorough treatment of such generalizations is beyond the scope of the present paper, we note that our results can be used to establish potentially useful lower bounds.

We illustrate such a bound for the case $K = 2$ of the example above. Specifically, consider a superposition dirty paper coding strategy analogous to that in the proof of the lower bound in Theorem 1, whereby we decompose the interferences according to (7). In this case, we have that (8) specializes to

$$\begin{aligned} S^n &= \beta_S\, S_0^n \\ V^n &= \beta_V\, S_0^n, \end{aligned} \tag{19}$$

where

$$\begin{aligned} \beta_S &= (\beta_1 + \beta_2)/2 \\ \beta_V &= (\beta_1 - \beta_2)/2. \end{aligned} \tag{20}$$

When we turn to implement step 2) of the encoding in the proof of the lower bound of Theorem 1, in which S^n is treated as inteference and V^n as noise, the results of [1] cannot be directly applied since the interferences S^n and V^n in (19) are correlated. However, the stronger version of Costa's results in [3] establishes that provided there is common randomness at the transmitter and receivers, any correlation between interference and noise can be ignored in encoding and does not affect the (worst case) capacity.[3]

Thus, we obtain the following lower bound on capacity of our example multicast channel with correlated interferences:

$$C_\beta \geq \max_{\{(P_1,P_2): P_1 \geq 0, P_2 \geq 0, P_1+P_2 \leq P\}} R_\beta(P_1, P_2), \tag{21a}$$

where

$$R_\beta(P_1, P_2) = \frac{1}{2}\log\left(1 + \frac{P_1}{1 + (\beta_1 - \beta_2)^2 Q/4 + P_2}\right) + \frac{1}{4}\log\left(1 + P_2\right). \tag{21b}$$

While the capacity of this channel remains unknown, we note that the rate (21) is nontrivial. Indeed, it can be quite large even when the INRs $\beta_1 Q$ and $\beta_2 Q$ are each large, provided that $|\beta_1 - \beta_2|$ is small.

V. Concluding Remarks

The capacity of the common-message Gaussian broadcast channel with transmitter side information is bounded. Our results establish that in contrast to the single-user channel, the lack of side information at receivers strongly limits capacity.

References

[1] M. H. M. Costa, "Writing on dirty paper," *IEEE Trans. Inform. Theory*, vol. 29, no. 3, pp. 439–441, May 1983.

[2] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York, NY: Wiley, 1991.

[3] U. Erez, S. Shamai (Shitz), and R. Zamir, "Capacity and lattice strategies for cancelling known interference," *IEEE Trans. Inform. Theory*, vol. 51, no. 11, pp. 3820–3833, Nov. 2005.

[4] S. I. Gel'fand and M. S. Pinsker, "Coding for channel with random parameters," *Probl. Peredachi Inform. (Probl. Inform. Trans.)*, vol. 9, no. 1, pp. 19–31, 1980.

[5] C. Heegard and A. E. Gamal, "On the capacity of computer memory with defects," *IEEE Trans. Inform. Theory*, vol. 29, pp. 731–739, Sep. 1983.

[6] A. Khisti, "Coding techniques for multicasting," Master's thesis, M.I.T, Cambridge, MA, 2004, http://web.mit.edu/khisti/www/SMThesis.pdf.

[7] A. Khisti, U. Erez, A. Lapidoth, and G. Wornell, "Carbon Copying onto Dirty Paper," *IEEE Trans. Inform. Theory*, submitted, November 2005, also available, http://de.arxiv.org/abs/cs.IT/0511095.

[8] A. Khisti, U. Erez, and G. Wornell, "Writing on two pieces of dirty paper at once," in *Proc. Int. Symp. Inform. Theory*, June 2004.

[9] C. E. Shannon, "Channels with side information at the transmitter," *IBM J. Res. Dev.*, vol. 2, pp. 289–293, Oct. 1958.

[10] Y. Steinberg and S. Shamai (Shitz), "Achievable rates for the broadcast channel with states known at the transmitter," in *Proc. Int. Symp. Inform. Theory*, 2005.

[3]The common randomness enables the channel input to be statistically independent of the effective noise.