Writing on dirty paper (Corresp.)

Article *in* IEEE Transactions on Information Theory · June 1983

DOI: 10.1109/TIT.1983.1056659 · Source: IEEE Xplore

1 author:

Max Costa
University of Campinas
47 PUBLICATIONS **4,634** CITATIONS

Lemma 1 we obtain the following codes:

Parameters of C	G' Belongs to
(271, 8, 134)	$\mathcal{G}(7, 6, 5, 4, 2)$
(274, 8, 136)	$\mathcal{G}(7, 6, 5, 4)$
(279, 8, 138)	$\mathcal{G}(7, 6, 5, 3, 2)$
(282, 8, 140)	$\mathcal{G}(7, 6, 5, 3)$
(286, 8, 142)	$\mathcal{G}(7, 6, 5, 2)$
(289, 8, 144)	$\mathcal{G}(7, 6, 5)$.

Finally, shortening each of these codes concludes the construction of $(g(8, d), 8, d)$ codes for all $d \in [131, 144]$.

VI. Conclusion

The purpose of this paper has been to show how known codes meeting the Griesmer bound can be described as a subfamily of a much wider class of such codes. One problem that has not been considered here is to decide all possible parameters these codes may have. This is a topic for further research. However, codes which meet the Griesmer bound for $k = 8$ have been constructed in this paper for all parameters for which such codes exist.

References

[1] J. H. Griesmer, "A bound for error-correcting codes," *IBM J. Res. Develop.*, vol. 4, pp. 532–542, Nov. 1960.

[2] G. Solomon and J. J. Stiffler, "Algebraically punctured cyclic codes," *Inform. Contr.*, vol. 8, pp. 170–179, Apr. 1965.

[3] B. I. Belov, "A conjecture on the Griesmer bound," in *Optimization Methods and Their Applications*, (Russian), Sibirsk. Energet. Inst. Sibirsk. Otdel. Akad. Nauk SSSR, Irkutsk, 1974, pp. 100–106, 182.

[4] H. C. A. van Tilborg, "On the uniqueness resp. nonexistence of certain codes meeting the Griesmer bound," *Inform. Contr.*, vol. 44, pp. 16–35, Jan. 1980.

[5] T. Helleseth and H. C. A van Tilborg, "The classification of all (145, 7, 72) binary linear codes," Eindhoven Univ. Tech., T. H. Rep. 80-WSK-01, Apr. 1980.

[6] ——, "A new class of codes meeting the Griesmer bound," *IEEE Trans. Inform. Theory*, vol. IT-27, pp. 548–555, Sept. 1981.

[7] T. Helleseth, "A characterization of codes meeting the Griesmer bound," *Inform. Contr.*, vol. 50, pp. 128–159, Aug. 1981.

[8] P. G. Farrell, "Linear binary anticodes," *Electron. Lett.*, vol. 6, pp. 419–421, June 1970.

[9] P. G. Farrell, "An introduction to anticodes," in *Algebraic Coding Theory and Applications*, G. Longo, Ed., CISM Courses and Lectures No. 258. New York: Springer-Verlag, 1979.

[10] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*. Amsterdam: North-Holland, 1977.

[11] B. I. Belov, V. N. Logachev, and V. P. Sandimirov, "Construction of a class of linear binary codes achieving the Varshamov–Griesmer bound," *Probl. Inform. Transm.*, vol. 10, pp. 211–217, July–Sept. 1974.

Correspondence

Writing on Dirty Paper

MAX H. M. COSTA

***Abstract*—A channel with output $Y = X + S + Z$ is examined. The state $S \sim N(0, QI)$ and the noise $Z \sim N(0, NI)$ are multivariate Gaussian random variables (I is the identity matrix.). The input $X \in R^n$ satisfies the power constraint $(1/n)\sum_{i=1}^{n} X_i^2 \leqslant P$. If S is unknown to both transmitter and receiver then the capacity is $\frac{1}{2}\ln(1 + P/(N + Q))$ nats per channel use. However, if the state S is known to the encoder, the capacity is shown to be $C^* = \frac{1}{2}\ln(1 + P/N)$, independent of Q. This is also the capacity of a standard Gaussian channel with signal-to-noise power ratio P/N. Therefore, the state S does not affect the capacity of the channel, even though S is unknown to the receiver. It is shown that the optimal transmitter adapts its signal to the state S rather than attempting to cancel it.**

Manuscript received October 19, 1981; revised August 1, 1982. This work was supported in part by CAPES (Brazil) Grant 4290/77 and in part by NSF Grant ECS78-23334. This work was presented at the IEEE International Symposium on Information Theory, Santa Monica, CA, February 1981.

The author was with the Information Systems Laboratory, Department of Electrical Engineering, Stanford University, Stanford, CA 94305; he is now with the Instituto de Pesquisas Espaciais (INPE), São José dos Campos, SP12200, Brazil.

I. Introduction

Consider the communication problem depicted in Fig. 1.

We wish to send an index $W \in \{1, \cdots, M\}$ to the receiver in n uses of the channel. Here M is the greatest integer smaller than or equal to e^{nR} and R is the rate in nats per transmission. The state $\boldsymbol{S}$ of the channel for n transmissions is assumed to be a sequence of independent identically distributed (i.i.d.) $N(0, Q)$ random variables. The state $\boldsymbol{S}$ is known to the transmitter but not to the receiver. Based on W and $\boldsymbol{S}$, the encoder sends a codeword $\boldsymbol{X}$ which must satisfy the power constraint $(1/n)\sum_{i=1}^{n} X_i^2 \leqslant P$.

The channel output is given by $\boldsymbol{Y} = \boldsymbol{X} + \boldsymbol{S} + \boldsymbol{Z}$, where the channel noise $\boldsymbol{Z}$ is distributed according to $N(0, NI)$. Upon receipt of $\boldsymbol{Y}$ the decoder creates an estimate $\hat{W}$ of the index W. The probability of error P_e, is given by

$$P_e = \frac{1}{M}\sum_{k=1}^{M} \Pr\{\hat{W} \neq k | W = k\}. \tag{1}$$

As it is seen in this equation, the probability of error is defined under the assumption that the index W is uniformly distributed over $\{1, \cdots, M\}$.

This is the standard Gaussian channel with input power constraint P, where the encoder is informed of part of the Gaussian

0018-9448/83/0500-0439$01.00

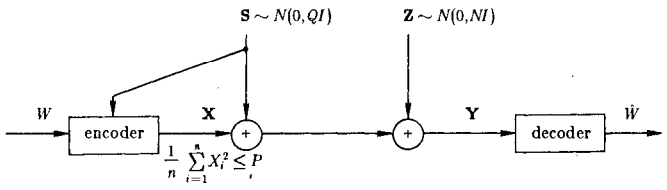

Fig. 1. Variation of Gaussian–Shannon channel.

additive noise sequence that will be added to his signal. Unfortunately, this information is not made available to the decoder, who will have to base his estimate $\hat{W}$ solely on the channel output Y.

Now imagine a sheet of paper covered with independent dirt spots of normally distributed intensity. In some sense, the problem of writing a message on this sheet of paper is analogous to that of sending information through the channel of Fig. 1. The writer knows the location and intensity of the dirt spots, but the reader cannot distinguish them from the ink marks applied by the writer. The title of this correspondence is based on this analogy, imprecise as it is.

There is an obvious encoding scheme to communicate over the channel of Fig. 1. If $P > Q$, the encoder may use part of his available power to cancel S. He can then use the remaining power $P - Q$ to send information at rate $C((P-Q)/N)$, where $C(x) = \frac{1}{2}\ln(1+x)$ is the capacity in nats/transmission of a Gaussian channel with signal-to-noise power ratio x. In general, using a fraction $0 \leqslant \alpha \leqslant \min\{1, Q/P\}$ of the transmitter power to partially cancel S yields a rate $C((1-\alpha)P/N + (\sqrt{Q} - \sqrt{\alpha P})^2)$. This scheme may be justified by one's temptation to reduce the problem to a previously solved one, but as we shall see, it is not an optimal encoding procedure. In fact the optimal encoding uses codewords in the direction of S. It looks at the space surrounding the vector S and chooses codewords that are compatible with the power constraint and far enough apart to be distinguishable when viewed from the channel output. In so doing, the encoder actually adapts its signal to the state S instead of trying to erase it.

In the next section, we show that all rates $R < C^* = C(P/N)$ are achievable. Then we use a simple argument to show that C^* is indeed the capacity of the model under investigation.

II. Capacity

Gel'fand and Pinsker [1] and El Gamal and Heegard [2] have shown that the capacity of a discrete memoryless channel with random state S known to the encoder is given by

$$C = \max_{p(u,x|s)} \{I(U;Y) - I(U;S)\}, \tag{2}$$

where the maximum is over all joint distributions of the form $p(s)p(u,x|s)p(y|x,s)$, where U is a finite alphabet auxiliary random variable.

Their random coding argument, assuming discrete, finite alphabets and unconstrained input, can be outlined as follows. First generate $\exp\{n(I(U;Y)-\epsilon)\}$ i.i.d. sequences U, according to the uniform distribution over the set of typical U.[1] Next, distribute these sequences uniformly over e^{nR} bins. For each sequence u let $i(u)$ be the index of the bin containing u. For encoding, given the state vector S and the message W, look in bin W for a sequence U such that (U, S) is jointly typical. Declare an error if no such U can be found. If the number of sequences in bin W is larger than $\exp\{n(I(U;S)+\delta)\}$, the probability of finding no such U decreases to zero exponentially as n increases. Next, choose X such that (X, U, S) is jointly typical and send it through the channel. At the decoder look for the unique sequence U such that (U, Y) is jointly typical. Declare an error if more than one or no such sequence exist. Then set the estimate $\hat{W}$ equal to the index of the bin containing the obtained sequence U. If $R < I(U;Y) - I(U;S) - \epsilon - \delta$, the probability of error averaged over all codes decreases exponentially to zero as $n \to \infty$. This shows the existence of a code that achieves rate R with arbitrarily small probability of error.

This result can be readily extended to memoryless channels with discrete time and continuous alphabets ([4, ch. 7]) by considering the supremum of $I(U_d;Y_p) - I(U_d;S_q)$ over all finite alphabet variables U_d and all partitions Y_p and S_q of the channel output and state alphabets.

The problem is reduced to that of finding an appropriate auxiliary variable U. We consider $U = X + \alpha S$, where X and S are independent random variables distributed according to $N(0, P)$ and $N(0, Q)$, respectively, and α is a parameter to be determined. Note that there could be a loss of generality in restricting attention to such U, but as we shall see, the derived answer is clearly optimal. Recalling that $Y = X + S + Z$ with Z distributed according to $N(0, N)$, the relevant mutual informations can be calculated to yield

$$\begin{aligned} I(U;Y) &= H(X+S+Z) - H(X+S+Z|X+\alpha S) \\ &= H(X+S+Z) + H(X+\alpha S) \\ &\quad - H(X+S+Z; X+\alpha S) \\ &= \frac{1}{2}\ln\left((2\pi e)^2(P+Q+N)(P+\alpha^2 Q)\right) \\ &\quad - \frac{1}{2}\ln\left((2\pi e)^2\left((P+Q+N)(P+\alpha^2 Q) - (P+\alpha Q)^2\right)\right) \\ &= \frac{1}{2}\ln\left(\frac{(P+Q+N)(P+\alpha^2 Q)}{PQ(1-\alpha)^2 + N(P+\alpha^2 Q)}\right) \end{aligned} \tag{3}$$

and

$$I(U;S) = \frac{1}{2}\ln\left(\frac{P+\alpha^2 Q}{P}\right). \tag{4}$$

Let

$$R(\alpha) = I(U;Y) - I(U;S). \tag{5}$$

Then

$$R(\alpha) = \frac{1}{2}\ln\left(\frac{P(P+Q+N)}{PQ(1-\alpha)^2 + N(P+\alpha^2 Q)}\right). \tag{6}$$

Graphs of $I(U;Y)$, $I(U;S)$, and $R(\alpha)$ as functions of α for $P = Q = N = 1$ are presented in Fig. 2.

Maximizing $R(\alpha)$ over α, we get

$$\max_{\alpha} R(\alpha) = R(\alpha^*) = \frac{1}{2}\ln\left(1 + \frac{P}{N}\right) = C^* \tag{7}$$

obtained for $\alpha^* = P/(P+N)$.

This is the desired answer. We know that the capacity of the channel cannot exceed $\max_{p(x|s)} I(X;Y|S)$, for this is the capacity when both encoder and decoder know the sequence S. This expression can be easily shown to equal C^*. But from (7), we achieve C^*. Thus the optimality of the present scheme is immediately established.

The above mutual informations evaluated for $U^* \equiv X + \alpha^* S$ are

$$I(U^*;Y) = \frac{1}{2}\ln\left(1 + \frac{P(P+Q+N)}{N(P+N)}\right) \tag{8}$$

and

$$I(U^*;S) = \frac{1}{2}\ln\left(1 + \frac{PQ}{(P+N)^2}\right). \tag{9}$$

[1] For a tutorial in the concepts of typicality and joint typicality, (see [3]).

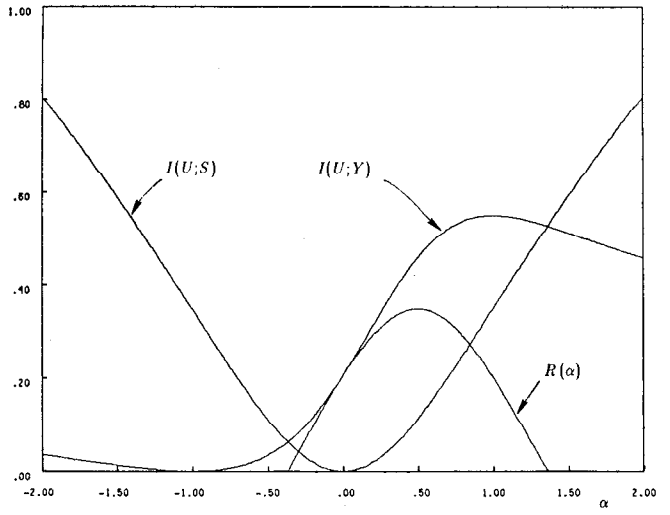

Fig. 2. Graphs of $I(U; Y)$, $I(U; S)$, and $R(\alpha)$ for $P = Q = N = 1$.

It is interesting to note that α^* as well as $R(\alpha^*)$ (the difference between these mutual informations) do not depend on Q.

In what follows we reexamine the code generation and the encoding–decoding procedures outlined before to show that we can send information at rate C^* with arbitrarily small probability of error and still satisfy the power constraint P. First, we generate $\exp\{n(I(U^*; Y) - \epsilon)\}$ sequences $\boldsymbol{U}$ ($\epsilon > 0$, arbitrarily small) with components independently drawn according to $N(0, P + \alpha^{*2}Q)$. Then, we place these sequences into $e^{nR} = \exp\{n(C^* - 2\epsilon)\}$ bins in such a way that each bin will contain the same number of sequences. The code book formed by the generated sequences and their assignments to the e^{nR} bins is known to both the encoder and the decoder. Given a state vector $\boldsymbol{S} = \boldsymbol{S}_0$ and a message $W = k$, the encoder looks for a jointly typical pair $(\boldsymbol{U}, \boldsymbol{S}_0)$ among the $\boldsymbol{U}$ in bin k. This is equivalent to looking for a sequence $\boldsymbol{U}$ such that

$$|(\boldsymbol{U} - \alpha^*\boldsymbol{S}_0)^t \boldsymbol{S}_0| \leq \delta \tag{10}$$

for some appropriately small δ. In other words, the encoder searches for the sequence $(\boldsymbol{U} - \alpha^*\boldsymbol{S}_0)$ which is nearly orthogonal to $\boldsymbol{S}_0$. The encoder declares an error if no such sequence is found. The number of sequences in each bin is $\exp\{n(I(U^*; S) + \epsilon)\}$. By arguments similar to those used in [5], [6] it can be shown that the probability of finding no suitable sequence $\boldsymbol{U}$ vanishes exponentially as $n \to \infty$. Call $\boldsymbol{U}_0$ the sequence obtained in this way. The encoder calculates $\boldsymbol{X}_0 = \boldsymbol{U}_0 - \alpha^*\boldsymbol{S}_0$. With high probability $\boldsymbol{X}_0$ will be typical, which is to say $\boldsymbol{X}_0$ will satisfy the power constraint $(1/n)\sum_{i=1}^{n}(X_{0i})^2 \leq P$. If not, the encoder declares an error.

Supposing no errors have occurred during the encoding procedure, the encoder sends $\boldsymbol{X}_0$ over the channel. Note that the set of X one might transmit is a continuum. At the other end, the decoder receives $\boldsymbol{Y} = \boldsymbol{Y}_0$ and then looks for a sequence $\boldsymbol{U}$ such that $(\boldsymbol{U}, \boldsymbol{Y}_0)$ is jointly typical. He declares an error if he can find more than one or no such sequence. With high probability the decoder will find only one such sequence and it will be equal to $\boldsymbol{U}_0$. Finally, the decoder sets his estimate $\hat{W}$ as the index of the bin containing this sequence. The probability of error averaged over the random choice of code decreases exponentially to zero as $n \to \infty$. This demonstrates the existence of a code satisfying the power constraint and achieving C^* with arbitrarily small probability of error.

III. Conclusion

It has been shown that the capacity of a channel with additive Gaussian noise and power constrained input is not affected if some extra i.i.d. noise sequence $\boldsymbol{S}$ is added to the output of the channel, as long as full knowledge of this extra noise sequence is given to the encoder. Rather than attempting to fight and cancel this extra noise, the optimal encoder adapts to it and uses it to his advantage, by choosing codewords in the direction of $\boldsymbol{S}$.

Interestingly, the set of signals that the optimal encoder might transmit is a continuum. Therefore, as we might expect, the optimal decoder will not have a good estimate of the transmitted signal nor will he have a good estimate of the sequence $\boldsymbol{S}$. He will decode a linear combination of the transmitted signal and the sequence $\boldsymbol{S}$.

The moral behind this result is: Do the best with what you have. This idea may also apply to other communication situations. A model in which the encoder can estimate the channel noise or where the encoder knows an interfering signal sent by a neighboring transmitter is a suitable candidate.

Acknowledgment

The problem was suggested to the author by Thomas Cover, who conjectured $C = C^*$. Thanks are due to him for his valuable help in the preparation of this paper. The author also thanks Abbas El Gamal, David Gluss, and Chris Heegard for many helpful discussions.

References

[1] S. I. Gel'fand and M. S. Pinsker, "Coding for channel with random parameters," *Problems of Control and Information Theory*, vol. 9, no. 1, pp. 19–31, 1980.

[2] A. El Gamal and C. Heegard, "On the capacity of computer memories with defects," submitted to *IEEE Trans. Inform. Theory*.

[3] A. El Gamal and T. M. Cover, "Multiple user information theory," in *Proc. IEEE*, vol. 68, no. 12, pp. 1466–1483, Dec. 1980.

[4] R. G. Gallager, *Information Theory and Reliable Communication*. New York: Wiley, 1968, ch. 7.

[5] A. El Gamal and E. van der Meulen, "A proof of Marton's coding theorem for the discrete memoryless broadcast channel," *IEEE Trans. Inform. Theory*, vol. IT-27, pp. 120–122, Jan. 1981.

[6] A. El Gamal and T. M. Cover, "Achievable rates for multiple descriptions," *IEEE Trans. Inform. Theory*, vol. IT-28, no. 6, pp. 851–857, November 1982.

The Discrete Memoryless Multiple Access Channel with Partially Cooperating Encoders

FRANS M. J. WILLEMS, STUDENT MEMBER, IEEE

Abstract—**We introduce the communication situation in which the encoders of a multiple access channel are partially cooperating. These encoders are connected by communication links with finite capacities, which permit both encoders to communicate with each other. First we give a general definition of such a communication process (conference). Then, by proving a converse and giving an achievability proof, we establish the capacity region of the multiple access channel with partially cooperating encoders. It turns out that the optimal conference is very simple.**

Manuscript received October 14, 1981; revised August 13, 1982. The research for this paper was supported by Onderzoeksfonds K. U. Leuven, Project OT/VI/17. This work was presented at the IEEE International Symposium on Information Theory, Les Arcs, France, June 21–25, 1982.

The author was with the Department Wiskunde, Katholieke Universiteit Leuven, Belgium. He is now with the Department of Electrical Engineering, Eindhoven University of Technology, Den Dolech 2, P.O. Box 513, 5600 MB Eindhoven, The Netherlands.

0018-9448/83/0500-0441$01.00 ©1983 IEEE