AD-A077 588 MASSACHUSETTS INST OF TECH LEXINGTON LINCOLN LAB F/G 17/2
A NOTE ON 'ALMOST NOISELESS' CHANNELS.(U)
OCT 79 B REIFFEN F19628-80-C-0002
UNCLASSIFIED TN-1979-66 ESD-TR-79-260 NL

1 OF 1
AD
A077588

END
DATE
FILMED
1-80
DDC

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
LINCOLN LABORATORY

A NOTE ON "ALMOST NOISELESS" CHANNELS

B. REIFFEN
Division 6

TECHNICAL NOTE 1979-66

4 OCTOBER 1979

LEXINGTON MASSACHUSETTS

ABSTRACT

An "almost noiseless" channel is defined. The definition corresponds to many physical channels operating at high signal-to-noise ratio. Attention is focused on the decrement in capacity and computation cutoff rate as the noise increases from zero. It is shown that the decrement in R_{comp} is dominated by a term proportional to the square root of the noisiest transition probability, α, to a nominally noiseless output. In contrast, the decrement in C in many cases of interest is dominated by a term proportional to $-\alpha \ln \alpha$. Accordingly the decrement in R_{comp} can significantly exceed the decrement in C.

This paper was presented at the 1979 IEEE International Symposium on Information Theory held in Grignano, Italy on 25-29 June 1979.

I. INTRODUCTION

This paper addresses the subject of memoryless channels which are "almost noiseless," i.e., channels where errors, suitably defined, are infrequent. More specifically, the capacity and R_{comp} behavior of these channels are addressed in terms of the noise characteristics.

This obscure theme is motivated by some practical problems in the engineering design of satellite communications systems. The context is given in Fig. 1. The satellite is not the familiar frequency translating repeater. Rather it is a "signal processing satellite" in which one or more uplinks are demodulated into bits which are then modulated onto a downlink carrier and beamed towards the receive terminal.

The motivations for performing this on-board signal processing are several: i) to avoid transmitting uplink noise or interference on the downlink as would be the case with a frequency translating repeater, and ii) to permit the decoupling of uplink and downlink modulation formats to better match hardware constraints. For example, it may be desirable to operate several individual uplinks FDMA and TDM on a single downlink carrier.

This situation is modelled as shown in Fig. 2. Note that the output of the uplink channel is directly connected to the input of the downlink channel without benefit of any decoding/reencoding operation. This constraint is associated with the desire to avoid the satellite complexity implied by any substantive on-board decoding. System error control is accomplished via end-to-end coding, i.e., encoding takes place at the transmit terminal and decoding takes place at the receive terminal.

The engineering design of such a system entails a number of complex trade-offs. A useful initial design approach is to specify the uplink modulation format and the end-to-end code assuming the downlink is noiseless. One would then perturb this first cut design to the extent that the downlink departs from the noiseless condition.

Hopefully this motivates an understanding of almost noiseless channels.

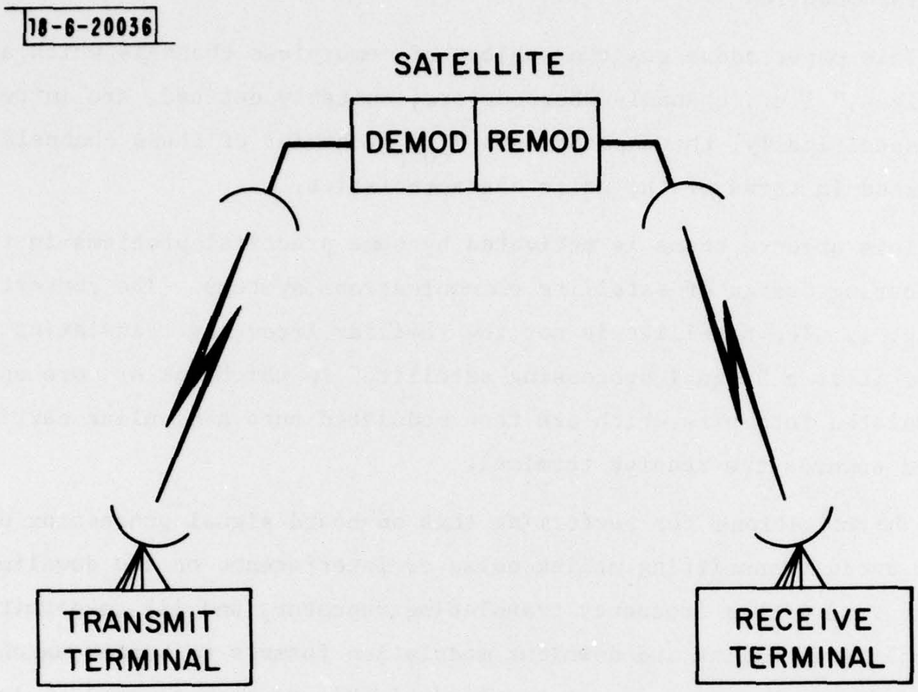


Fig. 1. Communication with a processing satellite.

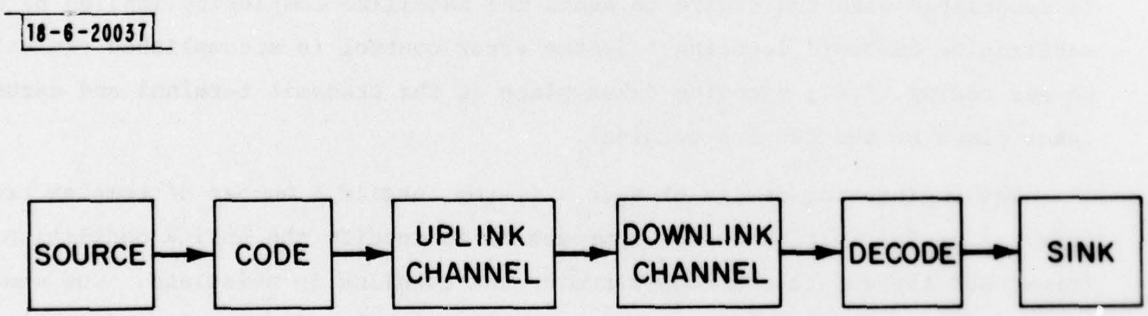


Fig. 2. Processing satellite model.

II. DEFINITIONS

A discrete memoryless channel is defined by

(a) an input alphabet, $X = \{x_i\}$, $i=1,\ldots,A$

(b) an output alphabet, $Y = \{y_j\}$, $j=1,\ldots,B$

(c) a transition probability matrix $\{q_{ij}\}$, where $\sum_i q_{ij} = 1$

The channel is defined as noiseless if and only if $q_{ij} = 0$ or 1 for all i,j. The input i and output j are called <u>connected</u> if $q_{ij} = 1$; i and j are <u>not connected</u> if $q_{ij} = 0$.

Suppose $b \leq B$ of the outputs are connected to some x_i. If any specific y_j is connected to more than one x_i, the information transmission capability of the channel is not impaired if these several x_i connected to the same y_j are treated as a single input. Thus, for noiseless channels, it is sufficient to let $A = b \leq B$.

The noiseless channel described above has capacity, C, and computation cutoff rate, R_{comp}, equal to

$$C = R_{comp} = \log A$$

Capacity is achieved when the A inputs are equally likely.

The noiseless channel described above is the limiting form of an idealization of a physical channel with "very large" signal-to-noise ratio. Actually, for finite signal-to-noise ratio, the channel will not be "noiseless". This paper addresses the question of how the characteristics of the channel change when "small" noise effects are present.

An <u>almost noiseless</u> channel is defined as one with a transition probability matrix $\{q_{ij}\}$ in the form

$$q_{ij} = \phi_{ij} + \alpha_{ij} \;, \quad (i=1,\ldots,A;\; j=1,\ldots,B;\; A \leq B)$$

where

$\{\phi_{ij}\}$ is the transition probability matrix of a noiseless channel,

and

$\{\alpha_{ij}\}$ is an AxB matrix with $\alpha_{ij} << 1$ for all i,j.

Since q_{ij} and ϕ_{ij} are each transition probability matrices, it is easy to show that

$$\sum_j \alpha_{ij} = 0 \quad , \; i = 1, \ldots, A \tag{1}$$

Since $\{q_{ij}\}$ is a probability matrix, $0 \leq q_{ij} \leq 1$ for all $\{i,j\}$. For x_i and y_j not connected via ϕ_{ij}, $\alpha_{ij} \geq 0$; for x_i and y_j connected via ϕ_{ij}, $\alpha_{ij} \leq 0$.

It will be convenient in the sequel to define several sets:

$$I = \{x_i, y_j | \phi_{ij} = 1\}$$

$$II = \{x_i, y_j | \phi_{kj} = 1 \text{ for some } k \neq i\}$$

$$III = \{x_i, y_j | \phi_{kj} = 0 \text{ for } k = 1, \ldots, A\}$$

$$Y_{NN} = \{y_j | \phi_{ij} = 1 \text{ for some } i\} \quad \text{(Nominally Noiseless outputs)}$$

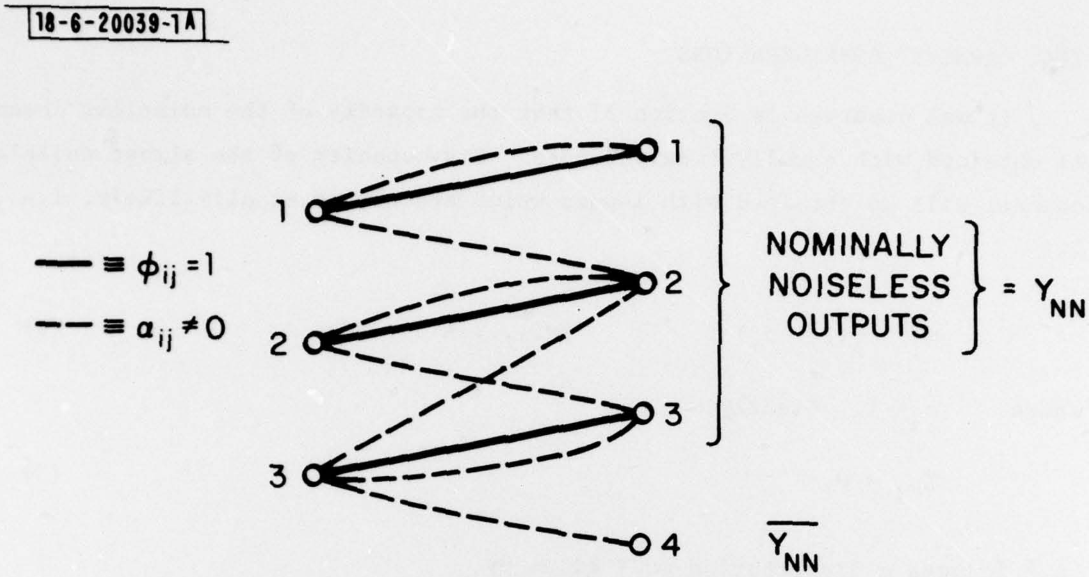
$$\overline{Y_{NN}} = \{y_j | \phi_{ij} = 0 \text{ for } i = 1, \ldots, A\} \quad \text{(Complement of } Y_{NN}\text{)}$$

Clearly,

$$I \cup II \cup III = XY$$

$$Y_{NN} \cup \overline{Y_{NN}} = Y$$

Figure 3 illustrates these sets with an example.



$(XY)_I$

X	Y
1	1
2	2
3	3

$(XY)_{II}$

X	Y
1	2
1	3
2	1
2	3
3	1
3	2

$(XY)_{III}$

X	Y
1	4
2	4
3	4

Fig. 3. Sets.

III. CAPACITY CONSIDERATIONS

It was observed in Section II that the capacity of the noiseless channel is obtained with equally likely inputs. The capacity of the almost noiseless channel will be obtained with inputs which are almost equally likely, i.e., let

$$p_i = \frac{1}{A}(1 + \sigma_i) \qquad i=1,\ldots,A \tag{2}$$

where $\sigma_i \ll 1$. Clearly,

$$\Sigma\sigma_i = 0. \tag{3}$$

$\{p_i\}$ induces a distribution on Y given by

$$r_j = \sum_i p_i q_{ij} = \frac{1}{A}\sum_i (1+\sigma_i)\ (\phi_{ij}+\alpha_{ij})$$
$$= \begin{cases} \frac{1}{A}[1+\sigma_j + \sum_i \alpha_{ij} + \sum_i \sigma_i\alpha_{ij}]\ , & y_j \subset Y_{NN} \\ \frac{1}{A}[\sum_i \alpha_{ij} + \sum_i \sigma_i\alpha_{ij}]\ , & y_j \subset \overline{Y_{NN}} \end{cases} \tag{4}$$

In Eq. (4), the convention has been adopted that $\phi_{ij} = 1$ if i=j, i.e., x_i is connected to y_j if and only if i=j.

The summation $\Sigma\sigma_i\alpha_{ij}$ in Eq. (4) involves the product of small numbers. Neglecting this second order term results in the following approximation

$$r_j = \begin{cases} \frac{1}{A}[1 + \sigma_j + \sum_i \alpha_{ij}]\ , & y_j \subset Y_{NN} \\ \frac{1}{A}\sum_i \alpha_{ij}\ , & y_j \subset \overline{Y_{NN}} \end{cases} \tag{5}$$

Now

$$I(X;Y) = H(Y) - H(Y|X)$$

$$-H(Y) = \sum_j r_j \log r_j = \sum_{Y_{NN}} \frac{1}{A}(1+\sigma_j + \sum_i \alpha_{ij})\log[\frac{1}{A}(1+\sigma_j + \sum_i \alpha_{ij})]$$

$$+ \sum_{\overline{Y_{NN}}} \frac{1}{A}(\sum_i \alpha_{ij})\log(\frac{1}{A} \sum_i \alpha_{ij}) \qquad (6)$$

Combining terms and utilizing Eqs. (1) and (3),

$$H(Y) = \log A - \frac{1}{A} \sum_{I+II} \alpha_{ij} - \frac{1}{A} \sum_{\overline{Y_{NN}}} (\sum_i \alpha_{ij})\log(\sum_i \alpha_{ij}) \qquad (7)$$

In going from Eq. (6) to (7), second order terms were neglected as before. Note that H(Y) is independent of first-order perturbations in the input probabilities.

$$-H(Y|X) = \sum_{i,j} p_i q_{ij} \log q_{ij}$$

$$-AH(Y|X) = \sum_{i,j} (1+\sigma_j)(\phi_{ij}+\alpha_{ij})\log(\phi_{ij}+\alpha_{ij})$$

$$= \sum_I (1+\sigma_i)(1+\alpha_{ij})\log(1+\alpha_{ij})$$

$$+ \sum_{II+III} (1+\sigma_i)\alpha_{ij}\log\alpha_{ij}$$

$$\approx \sum_I \alpha_{ij} + \frac{1}{A} \sum_{II+III} \alpha_{ij}\log\alpha_{ij} \qquad (8)$$

where the final approximation has neglected second order terms. Note that H(Y|X) is also independent of first order perturbations in the input probabilities. Combining Eqs. (7) and (8),

$$C = \log A - \frac{1}{A}\sum_{II}\alpha_{ij} - \frac{1}{A}\sum_{Y_{NN}}(\sum_i \alpha_{ij})\log(\sum_i \alpha_{ij})$$

$$+ \frac{1}{A}\sum_{II+III}\alpha_{ij}\log\alpha_{ij} \qquad (9)$$

where, in Eq. (9), H(Y)-H(Y|X) has been set equal to capacity since, to first order, the difference does not vary with small changes in the input probabilities.

$$\Delta C = \frac{1}{A}[\sum_{II}\alpha_{ij} + \sum_{Y_{NN}}(\sum_i \alpha_{ij})\log(\sum_i \alpha_{ij}) - \sum_{II+III}\alpha_{ij}\log\alpha_{ij}] \qquad (10)$$

Note that ΔC has two components, one from set II and one from set III. These components can be written as

$$\Delta C_{II} = -\frac{e}{A}\sum_{II}(\frac{\alpha_{ij}}{e})\log(\frac{\alpha_{ij}}{e}) \qquad (11)$$

$$\Delta C_{III} = \frac{1}{A}\sum_{Y_{NN}}(\sum_i \alpha_{ij})\,[\log(\sum_i \alpha_{ij}) - \sum_i \frac{\alpha_{ij}}{(\sum_i \alpha_{ij})}\log\alpha_{ij}] \qquad (12)$$

Since the log function is monotonically increasing, the bracketed expression in Eq. (12) is non-negative for all j. Accordingly, ΔC_{II} and ΔC_{III} are individually non-negative. Thus, it is seen that the capacity decrement is attributable to transitions to nominally noiseless outputs, and transitions to not nominally noiseless outputs, with the two kinds of transitions contributing somewhat differently.

For the BSC, the exact value of ΔC is $-p\log p - (1-p)\log(1-p)$ for $0 \leq p \leq 1$. The approximate value of ΔC obtained from Eq. (11)* is $-p\log p + p$. The ratio of these quantities is within 1% of unity for $p \leq 0.07$ and within 5% of unity for $p \leq 0.22$.

*Note that set III is empty for the BSC.

For the BEC, the exact value of ΔC is plog2 for $0 \leq p \leq 1$. The approximate value obtained from Eq. (12)* is also plog2. Thus, the approximation is exact for the BEC.

---

*Note that set II is empty for the BEC.

IV. R_{comp} CONSIDERATIONS

$$e^{-R_{comp}} = \sum_j [\sum_i p_i \sqrt{q_{ij}}]^2 = \sum_{i,j,k} p_i p_k \sqrt{q_{ij}} \sqrt{q_{kj}}$$

The summation on j will contain terms of the form

$$p_i p_k \sqrt{q_{ij}} \sqrt{q_{kj}} = \frac{1}{A^2} (1+\sigma_i) (1+\sigma_k)(\phi_{ij}+\alpha_{ij})^{1/2} (\phi_{kj}+\alpha_{kj})^{1/2}$$

$$= \begin{cases} \frac{1}{A^2} (1+\sigma_i)^2 (1+\alpha_{ij}) , & i = j = k \\ \frac{1}{A^2} (1+\sigma_i) (1+\sigma_k) (1+\alpha_{ij})^{1/2} \alpha_{kj}^{1/2}, & i = j \neq k \\ \frac{1}{A^2} (1+\sigma_i) (1+\sigma_k) \alpha_{ij}^{1/2} (1+\alpha_{kj})^{1/2}, & i \neq j = k \\ \frac{1}{A^2} (1+\sigma_i) (1+\sigma_k) \alpha_{ij}^{1/2} \alpha_{jk}^{1/2}, & i \neq j \neq k \end{cases}$$

Expanding, preserving only lowest order terms, and summing

$$A^2 e^{-R_{comp}} = A + 2\sum_{II} \sqrt{\alpha_{ij}} \qquad (13)$$

Note that to lowest order, R_{comp} is also independent of small variations in the input probabilities. From (13)

$$R_{comp} = \log \frac{A}{1 + \frac{2}{A} \sum_{II} \sqrt{\alpha_{ij}}}$$

$$= \log A - \log[1 + \frac{2}{A} \sum_{II} \sqrt{\alpha_{ij}}]$$

$$\approx \log A - \frac{2}{A} \sum_{II} \sqrt{\alpha_{ij}} \qquad (14)$$

$$\Delta R_{comp} = \frac{2}{A} \sum_{II} \sqrt{\alpha_{ij}} \quad (15)$$

Note that the decrement in R_{comp} is dominated by transitions to nominally noiseless outputs, Y_{NN}, i.e., the decrement will contain terms of order $\sqrt{\alpha_{ij}}$ for $y_j \subset Y_{NN}$. However, for $y_j \subset \overline{Y_{NN}}$, the decrement contains no terms of order $\sqrt{\alpha_{ij}}$.

For the BSC, the exact value of ΔR_{comp} is $\log(1 + 2\sqrt{p(1-p)}\,)$. The approximate value obtained from Eq. (15) is $2\sqrt{p}$. The ratio of these values is within 1% of unity for $p \leq 10^{-4}$ and within 5% of unity for $p \leq 3 \times 10^{-3}$.

For the BEC, the exact value of $\Delta R_{comp} = \log(1+2p) = 2p - 2p^2 + \ldots$. Note that the leading term is proportional to p. This is consistent with the approximation which predicts no contribution proportional to $\sqrt{p}$.*

---

*Recall that for the BEC, II is empty.

V. COMPARISON OF ΔC AND ΔR_{comp}

It has been noted that ΔC has two components: one contributed by Type II transitions and given by Eq. (11), and one contributed by Type III transitions and given by Eq. (12). On the other hand, ΔR_{comp} as expressed in Eq. (15) contains only Type II transitions. Forming the ratio of Eqs. (11) and (15),

$$\frac{\Delta C_{II}}{\Delta R_{comp}} = -\left(\frac{\sqrt{e}}{2}\right) \frac{\sum_{II} \beta_{ij} \log \beta_{ij}}{\sum_{II} \sqrt{\beta_{ij}}} \tag{16}$$

where

$$\beta_{ij} = \frac{\alpha_{ij}}{e} \tag{17}$$

Corresponding to every term in the numerator of Eq. (16) is a term in the denominator. The ratio of these terms is

$$\frac{\beta_{ij} \log \beta_{ij}}{\sqrt{\beta_{ij}}} = 2\sqrt{\beta_{ij}} \log \sqrt{\beta_{ij}} \tag{18}$$

The function $-w\log w$ is monotonic in w for $0 \leq w \leq \frac{1}{e}$. Thus over the range of interest,*

$$\frac{-\beta_{ij} \log \beta_{ij}}{\sqrt{\beta_{ij}}} \leq -2\sqrt{\beta_{max}} \log \sqrt{\beta_{max}} \tag{19}$$

where

$$\beta_{max} = \max_{II} \{\beta_{ij}\} = \frac{1}{e} \max_{II} \{\alpha_{ij}\} \tag{20}$$

Combining (16), (17), and (18)

$$\frac{\Delta C_{II}}{\Delta R_{comp}} \leq -\sqrt{e} \left(\frac{\alpha_{max}}{e}\right)^{1/2} \log\left(\frac{\alpha_{max}}{e}\right)^{1/2} \tag{21}$$

*Since $w = \sqrt{\beta} = \sqrt{\frac{\alpha}{e}}$, $w < \frac{1}{e}$ corresponds to $\alpha < \frac{1}{e}$.

The right hand side of (21) increases from zero as α_{max} increases, and will always be less than its maximum value of $e^{-1/2}$ for $\alpha_{max} < 1/e$. In other words, for an almost noiseless channel, the decrement in R_{comp} always exceeds ΔC_{II}.

For a channel where $Y = Y_{NN}$, $\Delta C = \Delta C_{II}$ and $\Delta R_{comp} > \Delta C$. For a channel with $\overline{Y_{NN}}$ not empty, ΔR_{comp} may be smaller or larger than ΔC depending on the relative magnitudes of $\{\alpha_{ij}\ II\}$ and $\{\alpha_{ij}\ III\}$.

Figure 4 displays the exact and approximate expressions for $\frac{\Delta C}{\Delta R_{comp}}$ for the BSC.

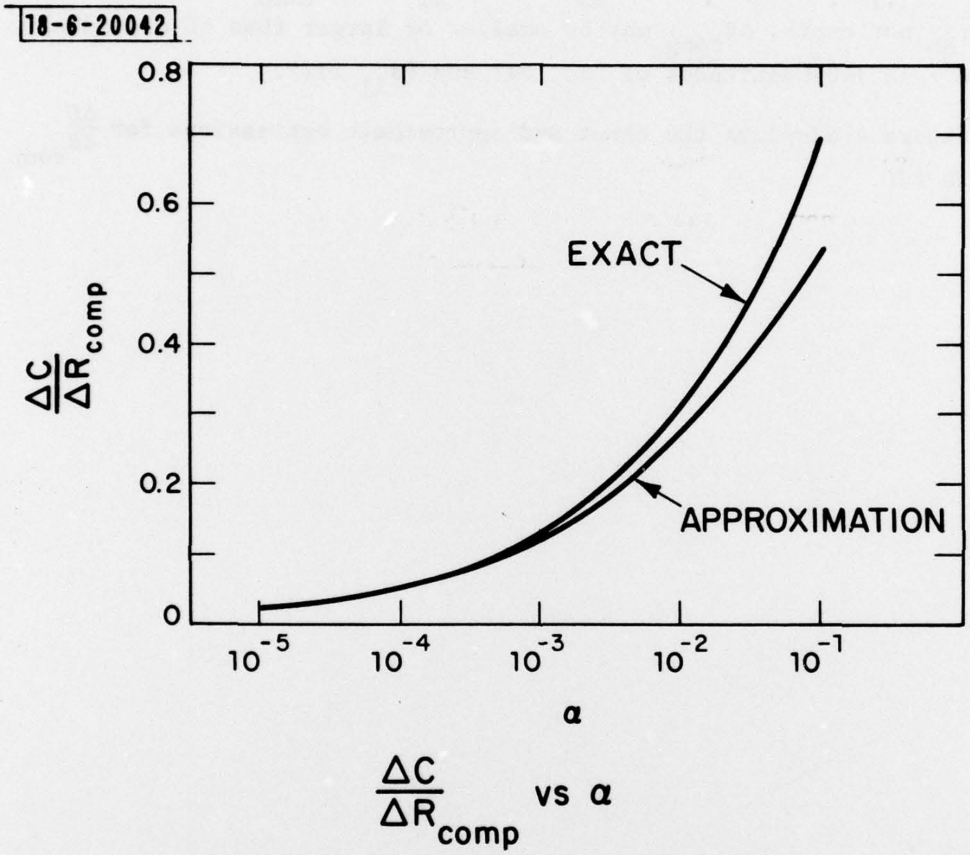


Fig. 4. $\Delta C/\Delta R_{comp}$ vs α for BSC.

VI. CONCLUSIONS

For "almost noiseless" discrete memoryless channels

a) C is achieved and R_{comp} is maximized with equally likely channel inputs

b) the decrement in C arises in part from terms of the form $(-\alpha\log\alpha)$, where $\alpha \ll 1$ is a noisy transition

c) the decrement in R_{comp} arises from terms of the form $\sqrt{\alpha}$

d) conclusions b) and c) imply that the decrement R_{comp} can significantly exceed the decrement in C.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER ESD-TR-79-260	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE *(and Subtitle)* A Note on 'Almost Noiseless' Channels		5. TYPE OF REPORT & PERIOD COVERED Technical Note
		6. PERFORMING ORG. REPORT NUMBER Technical Note 1979-66
7. AUTHOR(s) Barney Reiffen	TN-1979-66	8. CONTRACT OR GRANT NUMBER(s) F19628-80-C-0002
9. PERFORMING ORGANIZATION NAME AND ADDRESS Lincoln Laboratory, M.I.T. P.O. Box 73 Lexington, MA 02173		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS Program Element No. 65705F Project No. 649L
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Systems Command, USAF Andrews AFB Washington, DC 20331	20	12. REPORT DATE 4 Oct[redacted]979
		13. NUMBER OF PAGES 20
14. MONITORING AGENCY NAME & ADDRESS *(if different from Controlling Office)* Electronic Systems Division Hanscom AFB Bedford, MA 01731		15. SECURITY CLASS. *(of this report)* Unclassified
		15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT *(of this Report)*

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*

18. SUPPLEMENTARY NOTES

None

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*

satellite communications
almost noiseless memoryless channels
signal processing
coding

20. ABSTRACT *(Continue on reverse side if necessary and identify by block number)*

An "almost noiseless" channel is defined. The definition corresponds to many physical channels operating at high signal-to-noise ratio. Attention is focused on the decrement in capacity and computation cutoff rate as the noise increases from zero. It is shown that the decrement in R_{comp} is dominated by a term proportional to the square root of the noisiest transition probability, α, to a nominally noiseless output. In contrast, the decrement in C in many cases of interest is dominated by a term proportional to $-\alpha \ln \alpha$. Accordingly the decrement in R_{comp} can significantly exceed the decrement in C.

This paper was presented at the 1979 IEEE International Symposium on Information Theory held in Grignano, Italy on 25-29 June 1979.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

207 650