# XIV. PROCESSING AND TRANSMISSION OF INFORMATION[*]

Prof. R. G. Gallager
Prof. F. C. Hennie III
Prof. I. M. Jacobs
Prof. R. S. Kennedy
Prof. A. M. Manders
Prof. B. Reiffen
Prof. C. E. Shannon
Prof. I. C. Stiglitz
Prof. J. M. Wozencraft
E. R. Berlekamp
D. G. Botha
D. Chase

H. Dym
P. M. Ebert
D. D. Falconer
E. Ferretti
G. D. Forney, Jr.
N. Gramenopoulos
J. L. Holsinger
K. Ikushima
Y. Iwadare
G. P. McNamara
C. W. Niessen
J. D. Patterson
R. Pilc

G. W. Printz
L. H. Robinson
J. E. Savage
J. J. Sirota
J. R. Sklar
K. D. Snow
R. N. Spann
W. L. Stern
W. R. Sutherland
M. G. Taylor
H. L. Yudkin
H. P. Zeiger

## A. A LOWER BOUND ON THE DISTRIBUTION OF COMPUTATION FOR SEQUENTIAL DECODING — AN EXAMPLE

In the last quarterly report it was stated[1] that an overbound had been obtained on the distribution of the random variable of computation with the Wozencraft Sequential Decoding Algorithm. A lower bound has since been obtained for the distribution of this random variable. We shall present the lower bound by example, just as the upper bound was presented. Both examples deal with the mathematically amenable Binary Erasure Channel.

Both upper and lower bounds of the distribution decrease as some power of the distribution parameter, thereby establishing the algebraic character of the distribution of computation with Sequential Decoding.

### 1. The Algorithm

A stream of equiprobable and statistically independent binary digits from a source is encoded as follows: The first source digit selects a branch of $\ell$ binary digits from the first stage of a tree code (see Fig. XIV-1). The upper (lower) branch is selected if this first digit is 0 (1). Branches are selected at the second and later stages according to the same rule, and these second and later selections are dependent upon prior branch choices, that is, on prior source outputs.

The encoded source symbols are transmitted over a Binary Erasure Channel. This channel either erases a transmitted bit or passes it without alteration. The former event occurs with probability (1-C), where C is the capacity of this channel, and
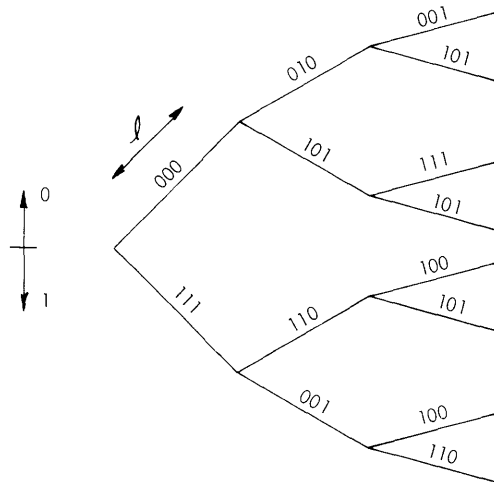
Fig. XIV-1. Tree code.

successive channel transitions are statistically independent. The transition diagram
for this channel is shown in Fig. XIV-2.

Corrupted channel digits arrive at the receiving terminal where they are processed.
The decoding algorithm is designed to determine the first-source output by using the
first S stages of the tree and, after this has been done, to decode the second-source
output by using the second through S + 1 stages of the tree, etc. We shall choose to
decode the first-source output with the aid of two identical machines, each of which
operates on a subset of the tree. (The tree is split into two subsets by the first digit.)
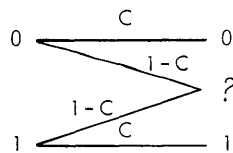Each machine compares the first branch of its subset with the first $\ell$ channel digits.



Fig. XIV-2. Binary Erasure Channel.

If any of these $\ell$ digits is unerased and in disagreement with a branch digit, this branch
could not have been transmitted and it becomes apparent that the unacceptable branch
is in the incorrect subset (the subset that does not contain the transmitted message).
The decoding process stops. Should both machines have acceptable first stages they
compare all branches at the second stage, comparing them one at a time with the second
set of $\ell$ channel digits, discarding unacceptable paths. The decoding process continues
either until a machine discards its subset, in which case it is decided that the other sub-
set is the correct subset and its corresponding source digit is produced as output, or
until one machine has looked at S stages, in which case the first-source symbol is ambig-
uous and the entire procedure terminates. Later digits are decoded in the same manner
by using the tree branching from previously decoded branches, provided that the process

has not terminated earlier. An ambiguity can be avoided (with high probability) by periodically resynchronizing, or by the use of feedback.

## 2. The Lower Bound

Let $X_1$, $X_2$ be the computations performed by the two machines. Then, a measure of the time spent in decoding the first digit is

$$X(n_t) = \min(X_1, X_2),$$

where $n_t = S\ell$ is the number of channel digits that are considered in decoding this first digit.

The probability that the total computation exceeds N, $P_r[X(n_t) > N]$, is certainly greater than the joint probability that the <u>total</u> computation exceeds N and the first $k\ell$ channel digits are erased, $P_r[X(n_t) > N,\ k\ell\ \text{erasures}]$. Similarly, this last probability is greater than the joint probability that the computation over the <u>first k branches</u>, $X(k\ell)$, exceeds N and the first $k\ell$ channel digits are erased, $P_r[X(k\ell) > N,\ k\ell\ \text{erasures}]$. Then

$$P_r[X(n_t) > N] > P_r[X(n_t) > N,\ k\ell\ \text{erasures}]$$

$$> P_r[X(k\ell) \geq N,\ k\ell\ \text{erasures}]$$

$$= P_r[X(k\ell) > N/k\ell\ \text{erasures}]\,(1-C)^{k\ell},$$

where $(1-C)^{k\ell}$ is the probability that the first $k\ell$ digits are erased. The conditional probability is either 1 or 0, since the computation over the first k branches, given that the first $k\ell$ digits are erased, is deterministic. Choose N so that the computation on the first k branches exceeds N, given the $k\ell$ erasures, so that this conditional probability has value 1. We then have

$$P_r[X(n_t) > N] \geq (1-C)^{k\ell},$$

in which we now choose the smallest k satisfying the necessary conditions. The required computation exceeds $2^{k-1}$, which is the number of branches in the last stage of one subset with k stages. If we choose k so that $2^{k-2} < N \leq 2^{k-1}$, we further lower-bound the distribution.

$$P_r[X(n_t) \geq N] > (1-C)^{2/R}\,(1/N)^{[\log_2(1/1-C)]/R}.$$

Here we have let $R = \dfrac{1}{\ell}$, the channel rate.

## 3. Conclusions

It has already been shown[1] that

$$P_r[X(n_t) \geq N] \leq 2\left\{\frac{3p}{\left[1 - 2^{-(1/p)[R_p/R-1]}\right]^3}\right\}^p \frac{1}{N^p}, \qquad R < R_p,$$

where the average is over the ensemble of all tree codes, and $R_p$ is interpreted geometrically from Fig. XIV-3. The exponent E(R) in this figure is the exponent on the probability of error obtained by "sphere packing."
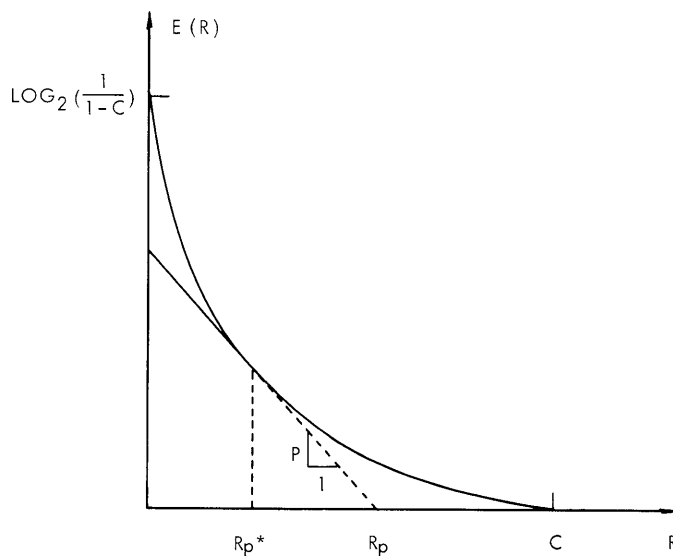


Fig. XIV-3. Construction for $R_p$.

Both upper and lower bounds decrease algebraically with N. The exponent on the lower bound is $\frac{E(0)}{R}$, which exceeds the exponent p on the upper bound for $R < R_p$.

J. E. Savage

References

1. J. E. Savage, A bound on the distribution of computation for sequential decoding — An example, Quarterly Progress Report No. 73, Research Laboratory of Electronics, M. I. T., April 15, 1964, pp. 155-164.