ESD-TR-66-337 MTP-40

ERROR STATISTICS OF A MULTIPLE LINK DATA TRANSMISSION CIRCUIT

FEBRUARY 1967

K. Brayer

Prepared for

DEVELOPMENT ENGINEERING DIVISION
DIRECTORATE OF AEROSPACE INSTRUMENTATION
ELECTRONIC SYSTEMS DIVISION
AIR FORCE SYSTEMS COMMAND
UNITED STATES AIR FORCE
L. G. Hanscom Field, Bedford, Massachusetts

Project 705B

Prepared by
THE MITRE CORPORATION
Bedford, Massachusetts

Contract AF19(628)-5165

ESD-TR-66-337

MTP-40

ERROR STATISTICS OF A MULTIPLE LINK DATA TRANSMISSION CIRCUIT

FEBRUARY 1967

K. Brayer

Prepared for

DEVELOPMENT ENGINEERING DIVISION
DIRECTORATE OF AEROSPACE INSTRUMENTATION
ELECTRONIC SYSTEMS DIVISION
AIR FORCE SYSTEMS COMMAND
UNITED STATES AIR FORCE
L. G. Hanscom Field, Bedford, Massachusetts

Project 705B

Distribution of this document is unlimited.

Prepared by

THE MITRE CORPORATION
Bedford, Massachusetts

Contract AF19(628)-5165

FOREWORD

This report was prepared by the Range Communications Planning and Technology Subdepartment of The MITRE Corporation, Bedford, Massachusetts, under Contract AF 19(628)-5165. The work was directed by the Development Engineering Division under the Directorate of Aerospace Instrumentation, Air Force Electronics Systems Division, Laurence G. Hanscom Field, Bedford, Massachusetts. Captain J. J. Centofanti served as the Air Force Project Monitor for this program, identifiable as ESD (ESSI) Task 5932.07, Range Digital Data Transmission Improvement.

REVIEW AND APPROVAL

Publication of this technical report does not constitute Air Force approval of the report's findings or conclusions. It is published only for the exchange and stimulation of ideas.

OTIS R. HILL, Colonel, USAF
Director of Aerospace Instrumentation

ABSTRACT

Error patterns have been measured on a combined troposcatter, microwave, wireline circuit of more than 6000 circuit miles. The error data consist of more than 6×10^8 transmitted bits at rates of 1200 and 2400 bits/sec. These error patterns are presented in a form suitable for the analysis of error control, modem improvement or channel modeling.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

SECTION I
INTRODUCTION

In the spring of 1966, data transmission tests were conducted by The MITRE Corporation on a U. S. Air Force data circuit which will be designated as OL-5/STC. The objective of these tests was to determine the characteristics of the error patterns which occur on this link.

The link comprises three types of transmission media: troposcatter, microwave, and wireline. The troposcatter and microwave are multiple-hop circuits and the wireline consists of numerous interconnected leased telephone wirelines. Measurements made on other parts of the circuit indicate that the troposcatter dominates in terms of error generation. Rixon Sebit 24B vestigial sideband AM modems were used at data rates of 1200 and 2400 bits/sec in a 3kc channel. The data collected consisted of 38 hours at 1200 bits/sec and 61 hours at 2400 bits/sec for a total of 1.67×10^8 and 5.20×10^8 bits respectively.

The error pattern data were obtained by one-way transmission of a 52-bit digital data test pattern which was compared with a locally generated test pattern at the receiving end. When there was agreement between the two, a logical zero was declared; when the bits were different, a logical one was declared. The zeros and ones were recorded on magnetic tape. These magnetic tapes were then processed through a tape-to-tape converter facility to produce an IBM compatible tape for processing in an IBM 7030 computer for determination of the statistics of these bit-by-bit error patterns.

SECTION II

DESCRIPTIVE STATISTICS

The data for the two transmission rates were processed for four types of statistics: the occurrence of consecutive errors, the occurrence of intervals between errors, the occurrence of bursts, and word error rate information. The extent of the data analyzed was as follows:

Data Rate	Total Bits	Average Error Rate
1200 b/s	1.67×10^{8}	1.39×10^{-4}
2400 b/s	5.20×10^{8}	3.95×10^{-4}

Figure 1 presents the relative frequency of occurrence of consecutive errors for the two data rates. The figure indicates that the relative occurrence of multiple errors is not a function of the data rate for these two rates. Further, the relative occurrence is certainly not of an independent random nature. If it were, over 98 percent of the errors would occur as single errors, while in fact only 70 percent occur as single errors. This lack of independence is further illustrated in Figure 2 where the cumulative frequency of occurrence of error-free gaps is presented for the actual data and the case of independent errors, respectively. All these plots confirm that this error pattern data is widely different from independent random errors at the same error rate. The cumulative distribution of error-free gaps for independent random errors is (in theory) described by the equation

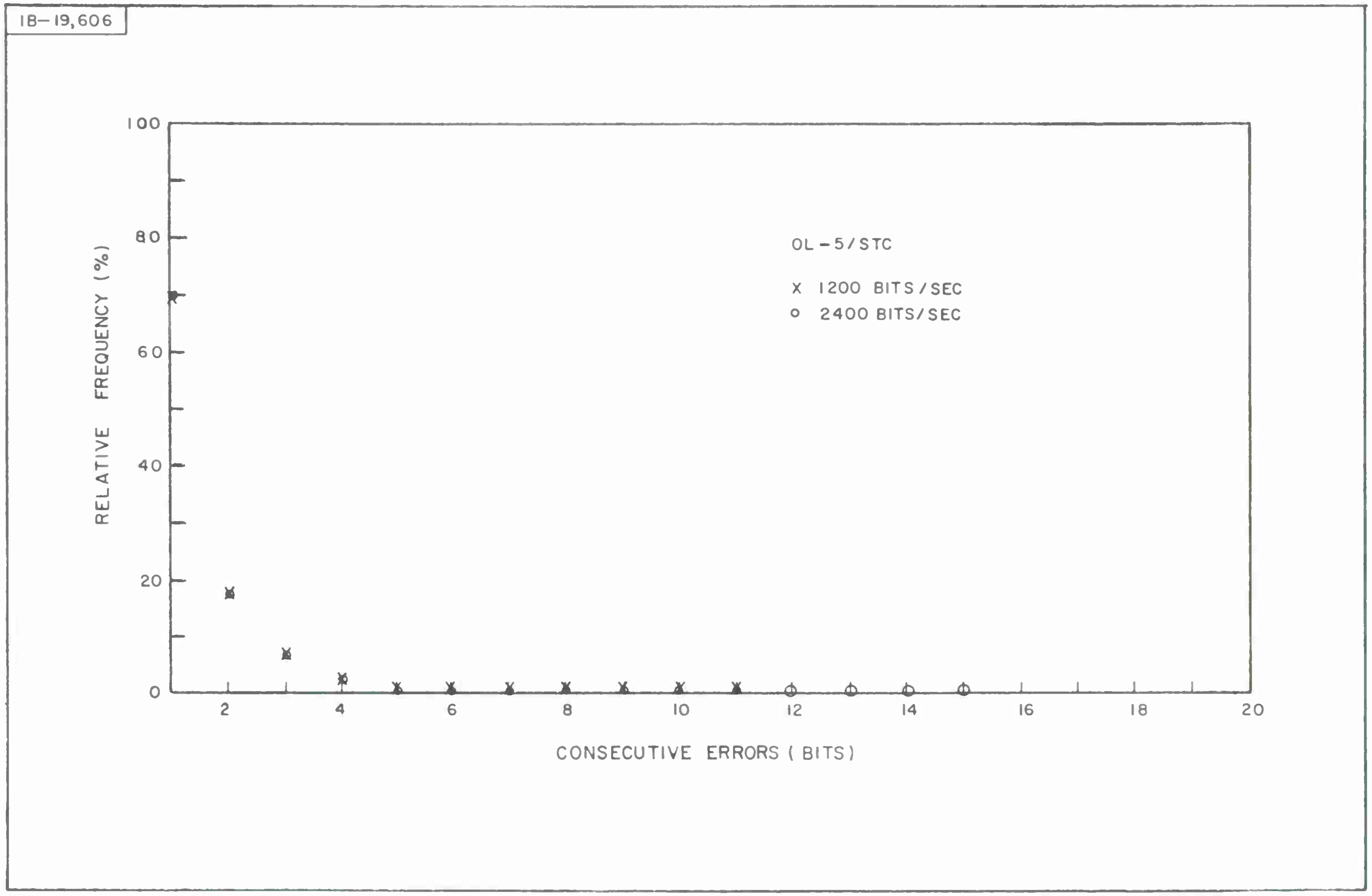

Figure 1. Frequency of Consecutive Error Occurrence

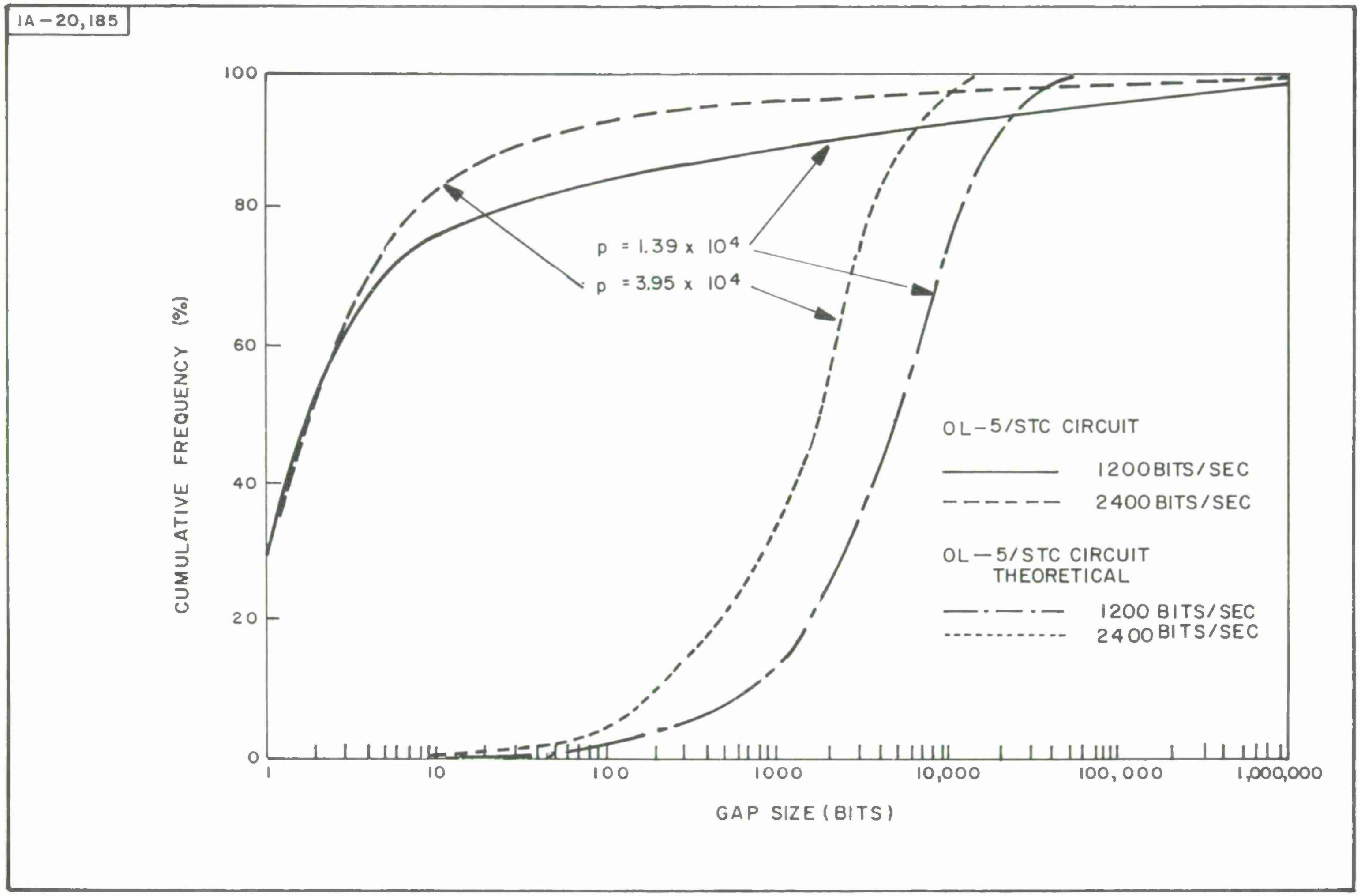

Figure 2. Distribution of Gaps Between Errors

$$P\left\{c^n e \mid e\right\} = \sum_{k=0}^{n} p(1-p)^k \qquad (1)$$

where

c represents a correct bit;

e represents an error bit;

n represents number of consecutive bits; and

p is the probability of bit error.

Figure 3 presents the message error rate (defined as the occurrence of at least one error in a message) as a function of message size for these two classes of data. Along with these curves, the corresponding values for independent error occurrence are presented. These values are derived from the relation

$$p_m = 1 - (1-p_e)^m \qquad (2)$$

where

p_m = probability of a message error;

p_e = probability of bit error; and

m = message length in bits.

The interesting thing to note in these curves is that the value of p_m for $m > 2000$ in the actual data is independent of error rate and signaling speed. This indicates that the errors are occurring in bursts, and, while for the overall data p_e at 2400 bits/sec is 2.85 times as large as p_e at 1200 bits/sec, the additional errors which are occurring at the higher signaling speed occur as increased burst density. Thus it is evident that

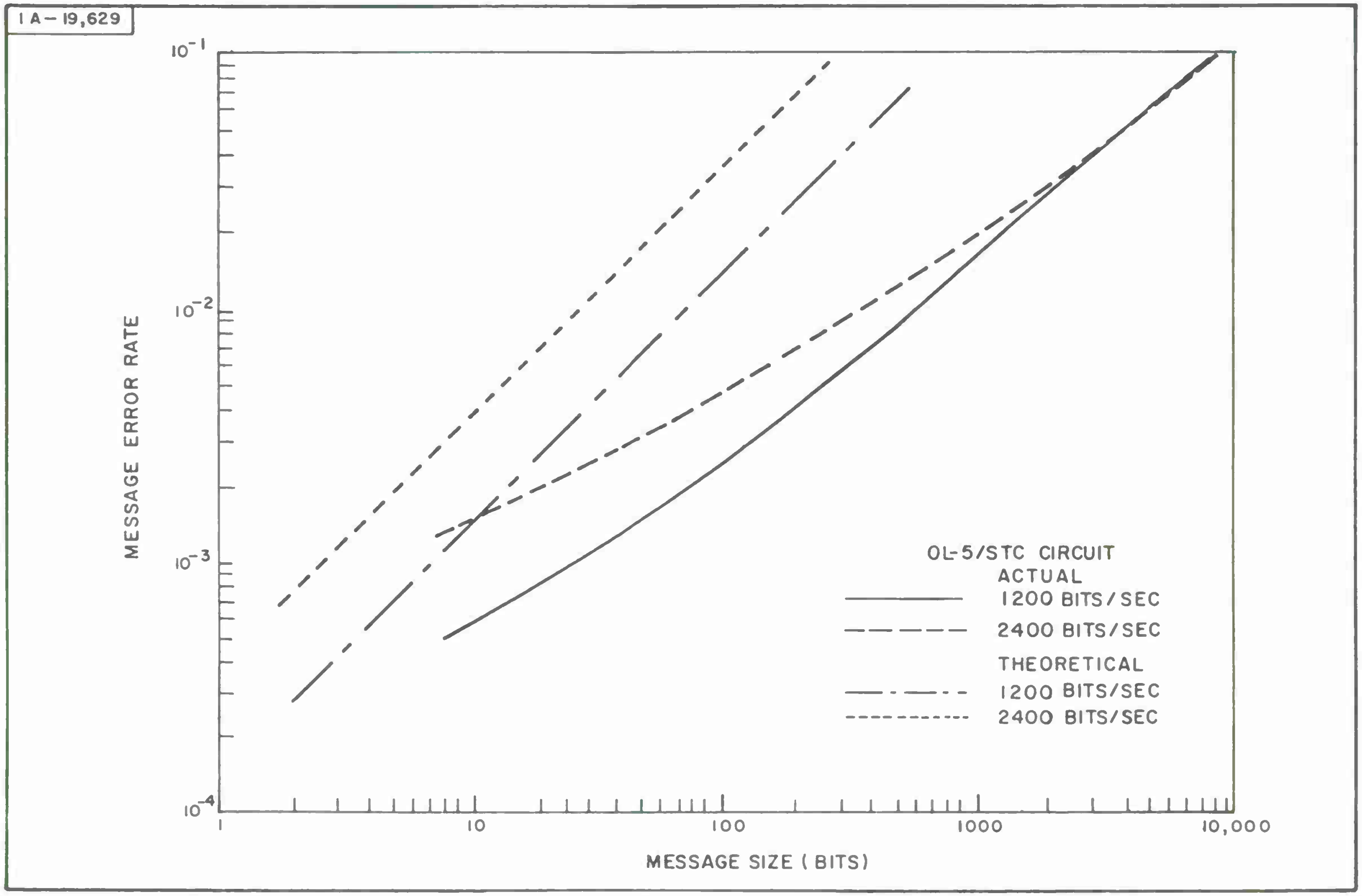

Figure 3. Message Error Rate vs Message Size

bursts occur as fades in time unrelated to the data rate (for the rates measured). This conclusion is confirmed by a visual observation of the error patterns.

SECTION III

DESCRIPTION OF BURST STATISTICS

The description of the error statistics presented thus far has made it clear that the errors fall in a non-random clumped distribution. To describe these error bursts and the intervals between bursts (guard space), an additional statistic has been developed.

DEFINITION OF BURST

A burst is defined as a region of the serial data stream where the following properties hold. A minimum number of errors, M_e, is contained in the region, and the minimum density of errors in the region is Δ. Both of these conditions must be satisfied for the chosen values of M_e and Δ for the region to be defined as a burst. The density of errors is defined as the ratio of bits in error to the total number of bits in the region.

The following properties hold for the burst. The burst always begins with a bit in error and ends with a bit in error. A burst may contain correct bits. Each burst is immediately preceded and followed by an interval in which the density of errors is less than Δ.

The burst probability density function is defined as the probability of occurrence of a burst of size N, where N is any positive integer. The burst size is measured in terms of the total number of bits in the burst. A separate burst probability density function may be determined for each pair of values of Δ and M_e.

The minimum number of errors in a burst has been chosen to be two for all the data included here. It was found that larger values of M_e would not change the values of mean burst length significantly. However, the intervals between the bursts were found to increase drastically so that little meaningful data could be obtained for the burst to consecutive interval ratio. When a value of one is selected for M_e, every error becomes a burst, and the requirement that a burst begin and end with an error is violated. Consequently, no meaningful data is obtained for this value of M_e.

DEFINITION OF INTERVAL

The interval is defined as the region of the serial data stream where the following properties hold. The minimum density of errors is less than Δ, and the region begins and ends in a correct bit. An interval may contain errors. An interval is always immediately preceded and followed by a burst. Thus, each and every bit in the data stream must lie in either a burst region or an interval region.

The interval probability density function is defined as the probability of occurrence of an interval of length L, where L is any positive integer. The interval probability density is a joint function of both Δ and M_e.

DESCRIPTIONS OF DISTRIBUTIONS

Figure 4 presents the distribution of observed burst lengths for the actual data. The burst lengths range from two (M_e) to over 1,000,000 bits. While 90 percent are less than 1000 bits long, the remaining bursts contain most of the errors. The densities of these bursts are displayed in Figure 5. This figure confirms the previous conclusions that at 2400 bits/sec, the additional errors serve to increase the density of the bursts. Further, while

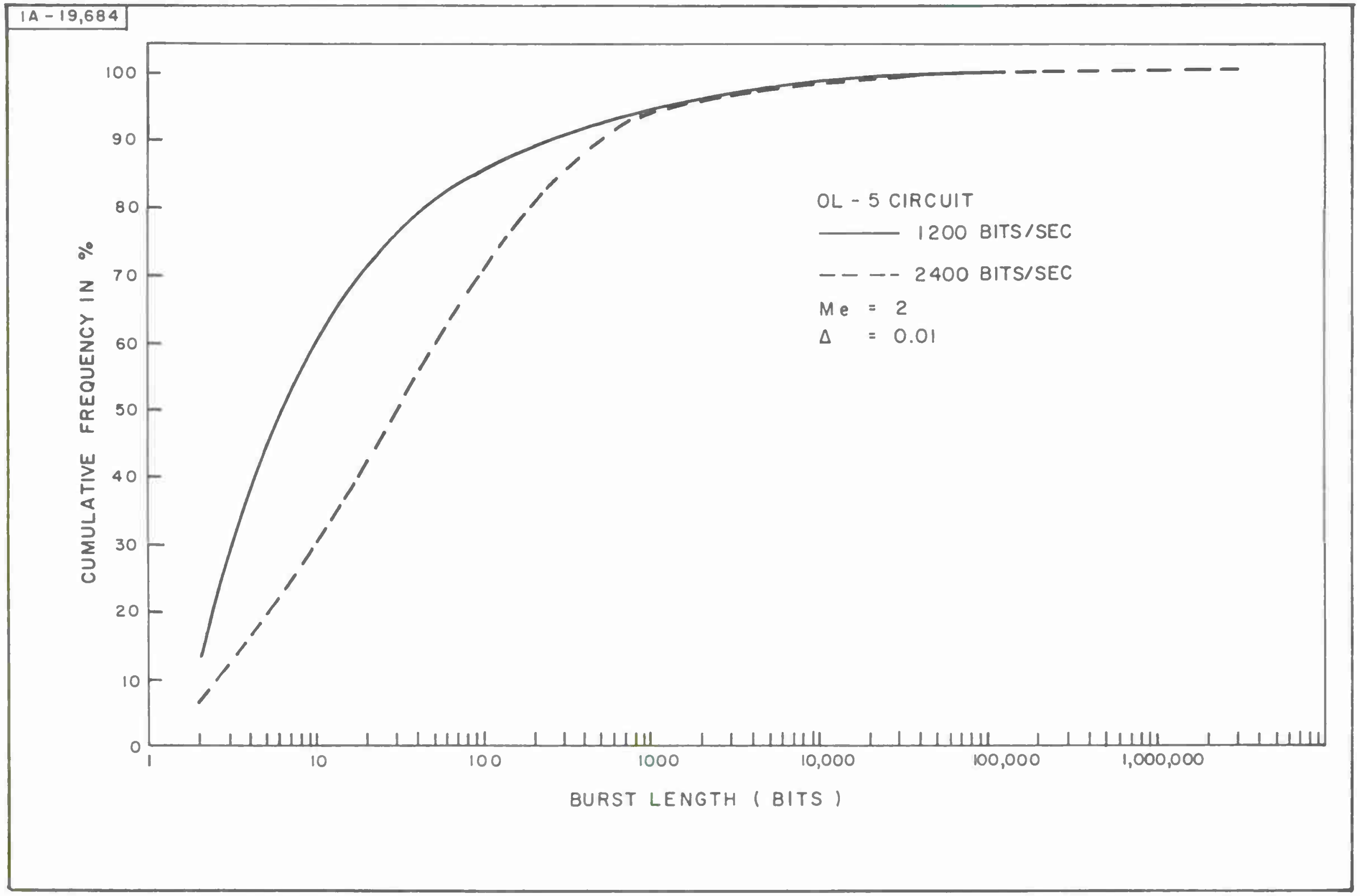

Figure 4. Distribution of Lengths of Bursts

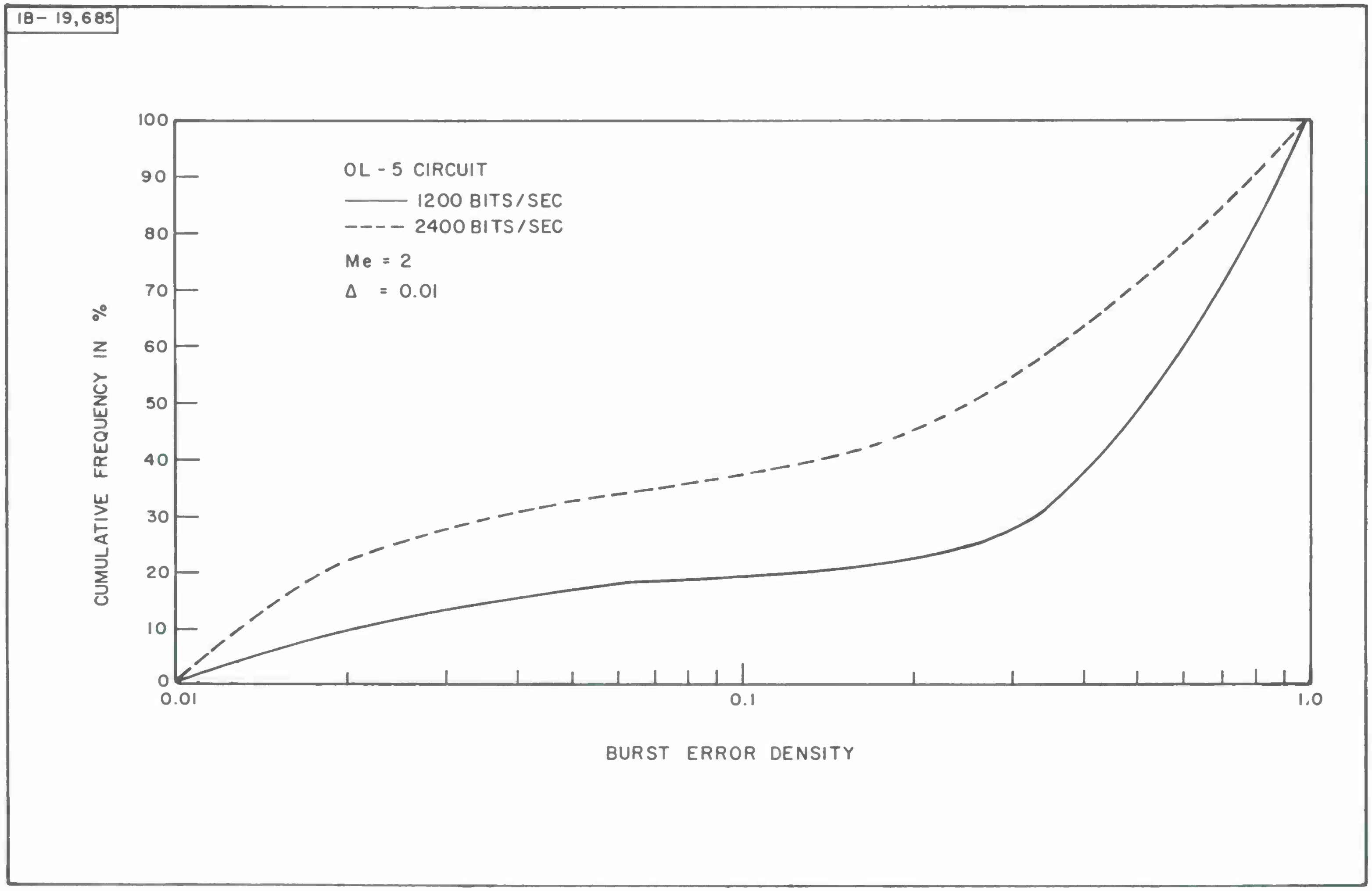

Figure 5. Distribution of Burst Error Densities

the minimum criteria on (Δ) is chosen at 0.01, over 50 percent of the bursts have densities greater than 0.1. A correlation study was performed on the burst lengths and their associated densities; and it was found that, while bursts less than 10 bits in length are generally 100 percent dense (consecutive errors) and bursts over 100,000 bits in length are near the minimum criteria, the other bursts range in density between the minimum and 75 percent, indicating no correlation between the length of bursts and densities. Figure 6 indicates the distribution of 1200 b/s data in this study. The value of Δ was chosen at 0.01 because only for this value of Δ are burst length and density independent of Δ. As a convenience in plotting the actual and theoretical curves, the values have been shifted up by one (since there is no zero on a log scale).

The distribution seen on these curves indicates that the errors fall into adjacent clusters of errors with long error-free intervals between clusters. There are no intermediate groupings of errors. In a theoretical, random error distribution, the intermediate groupings would predominate.

Figures 7 and 8 show the distribution of interval lengths and associated densities. The important facts to note here are that over 60 percent of the intervals are error free, and the remaining intervals have no more than one or two random errors within them.

If the length of the interval directly following every burst is divided by the burst length, a quantity known as the guard space ratio, which is useful in coder design, is obtained. Figure 9 indicates that almost 5 percent of the guard space ratios of this data link are insufficient (< 1) for direct code error correction, but that the remaining ratios are sufficient. Unfortunately, this 5 percent consists of the longest bursts.

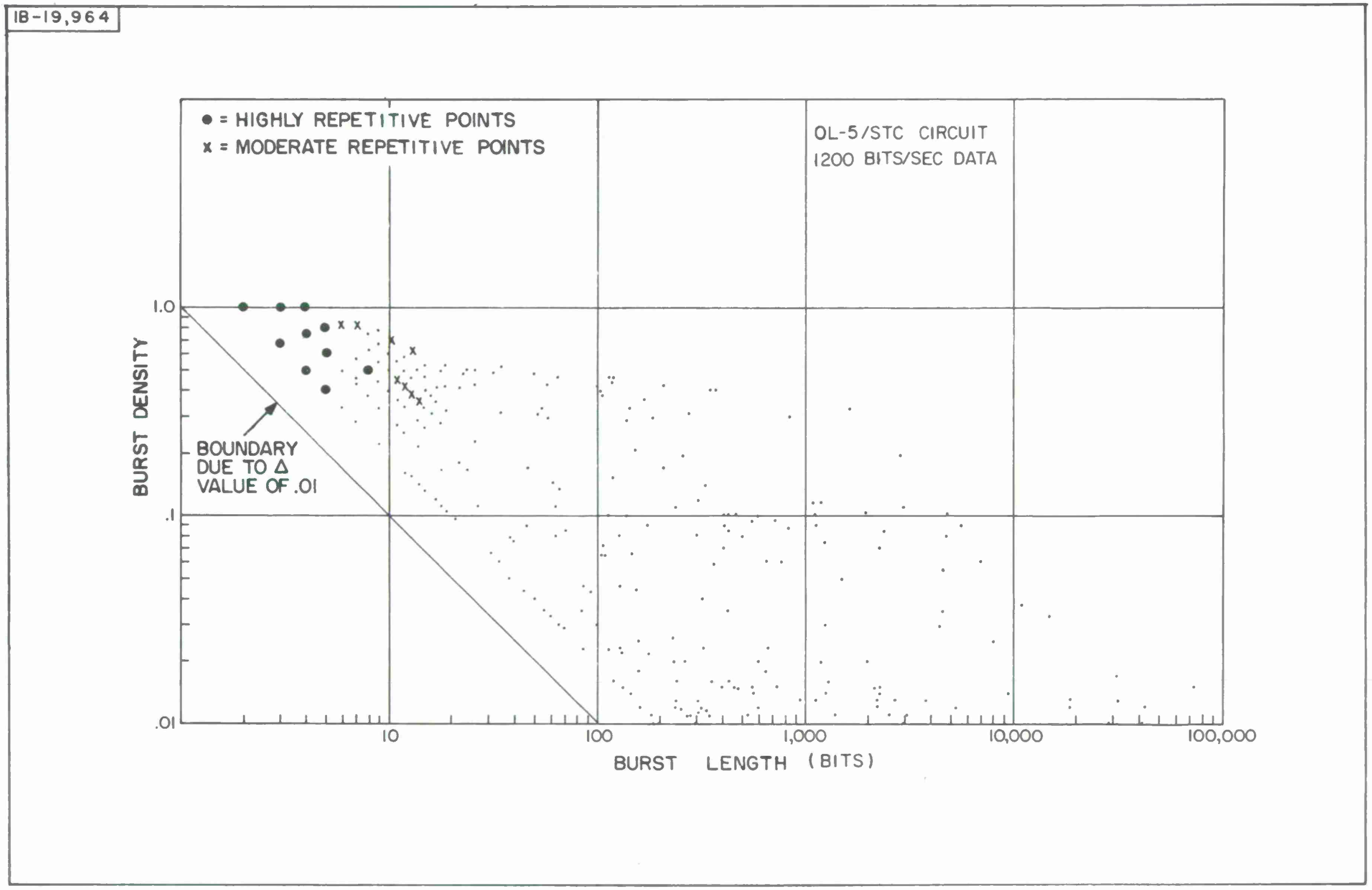

Figure 6. Relationship Between Burst Length and Burst Density

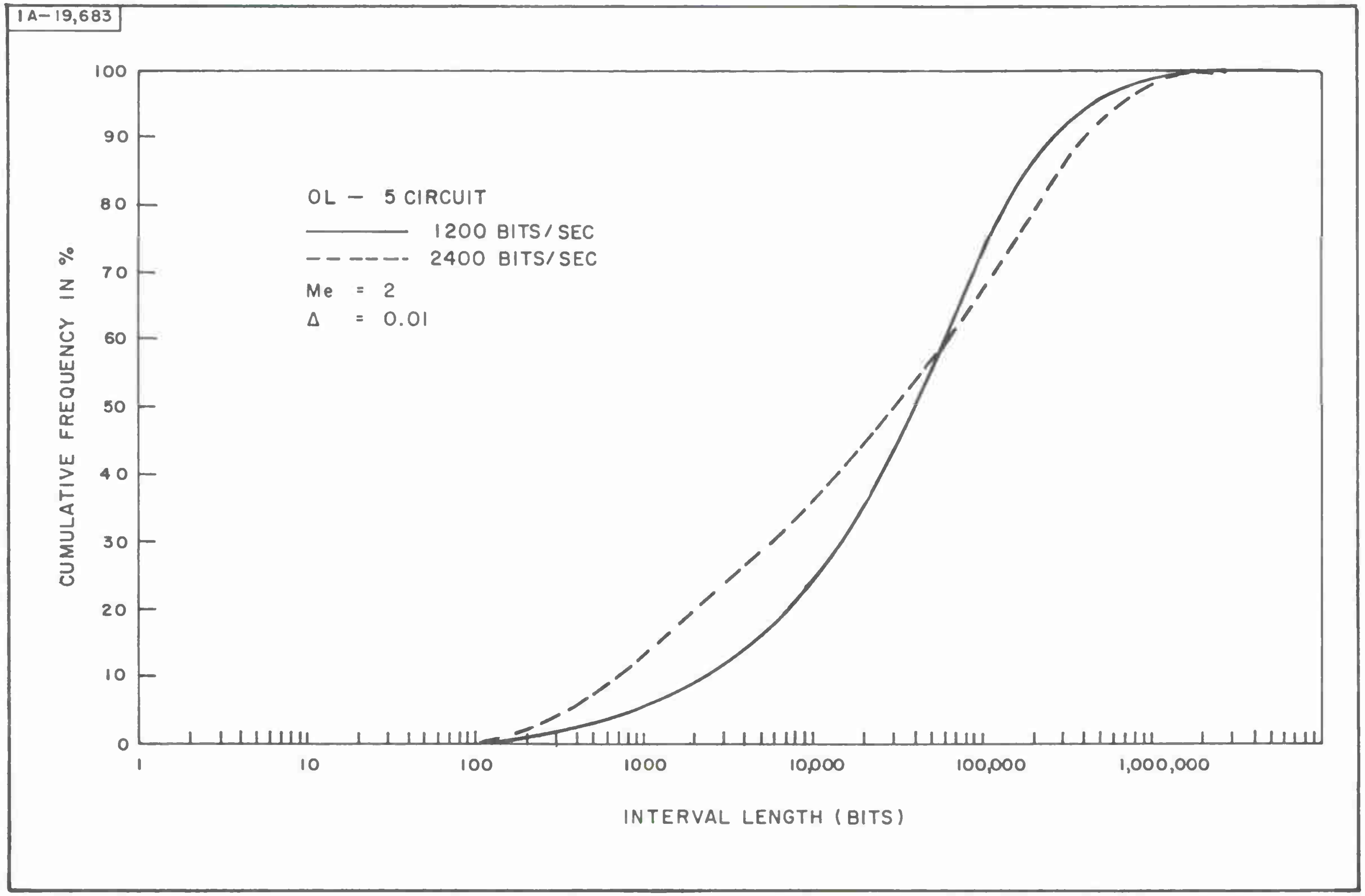

Figure 7. Distribution of Lengths of Intervals

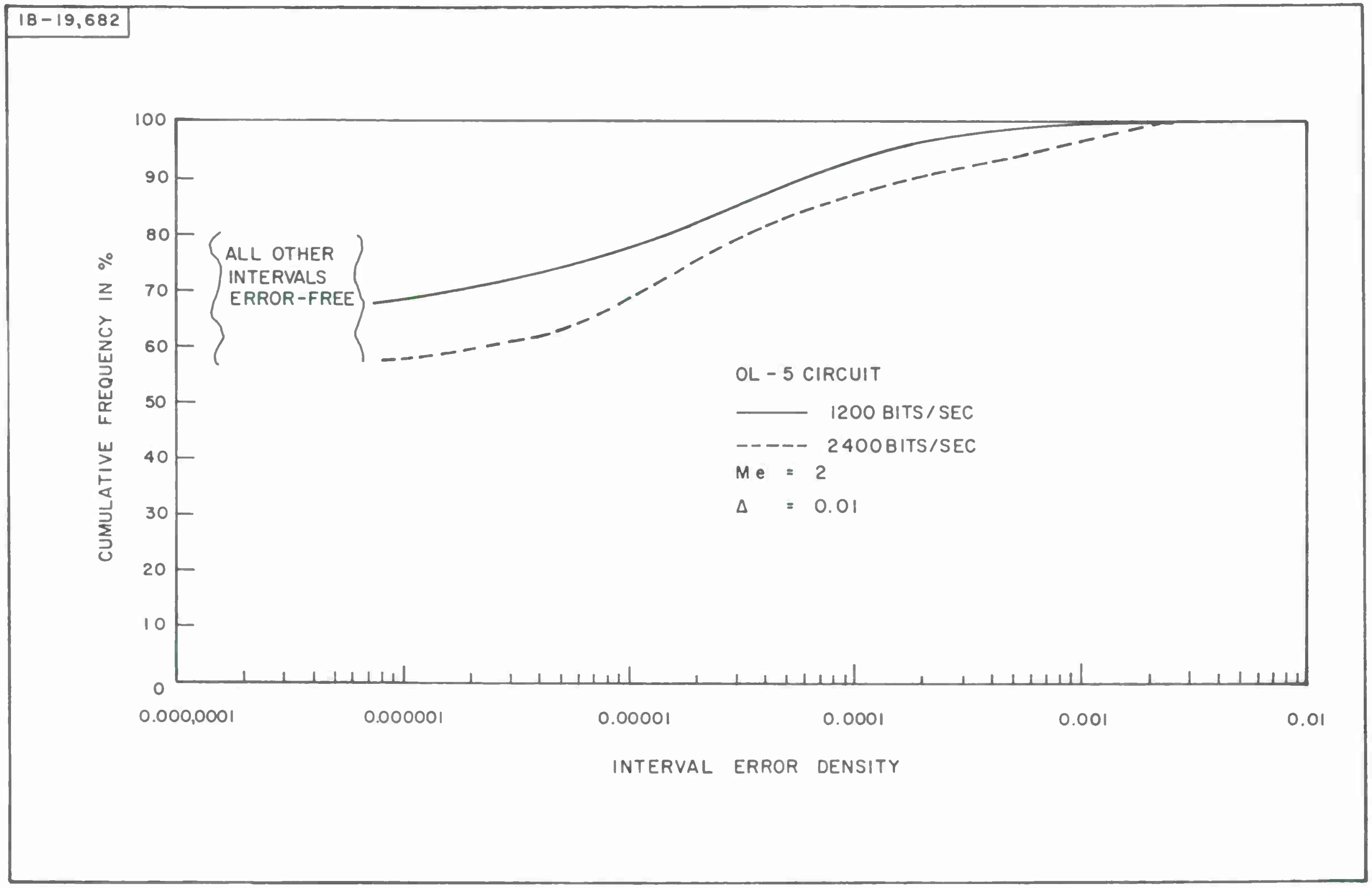

Figure 8. Distribution of Interval Error Densities

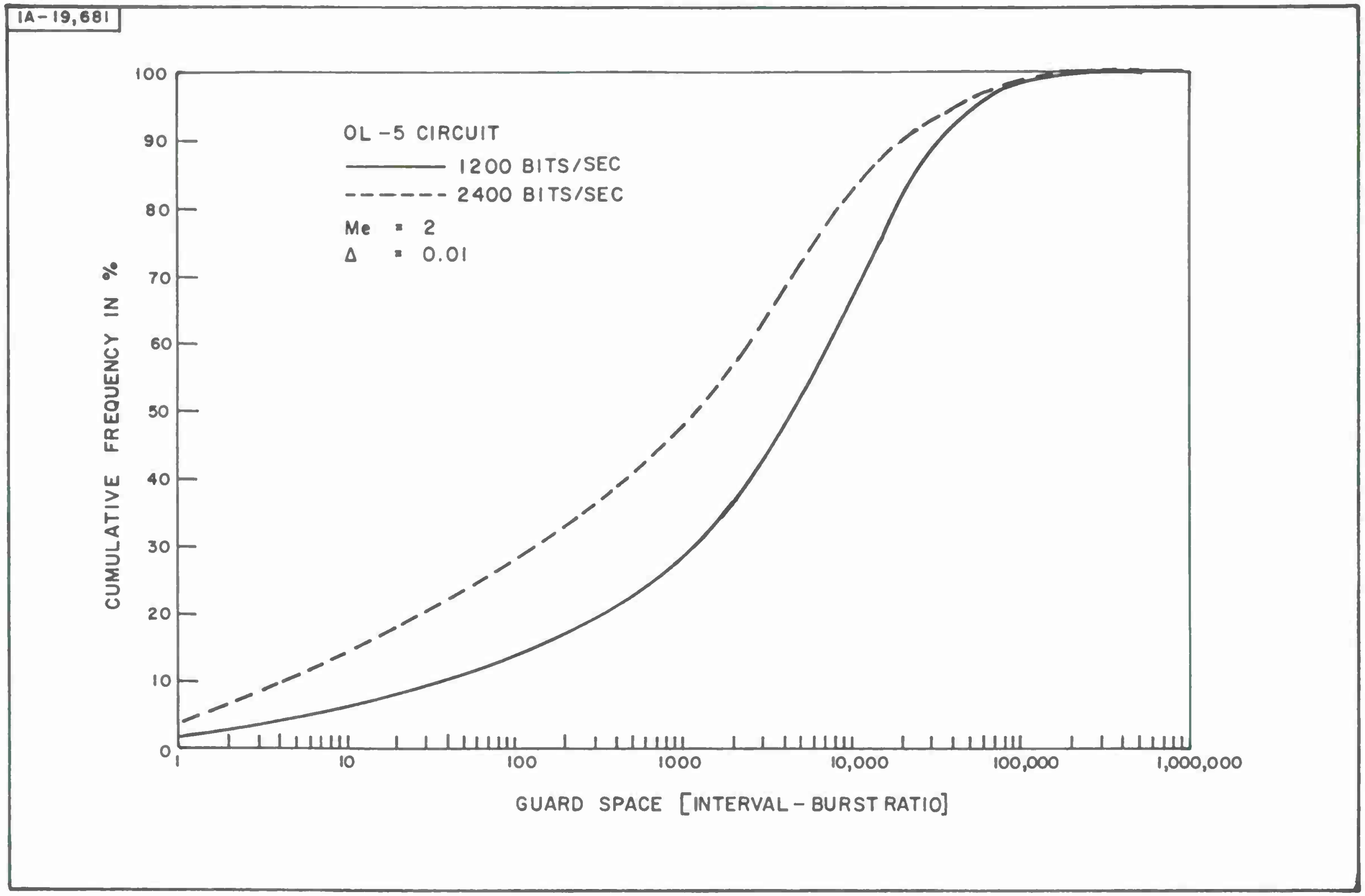

Figure 9. Guard Space Ratio Distribution

SECTION IV

CONCLUSIONS

The purpose of this paper is to present the error patterns which occur on a circuit predominated by a troposcatter path. As indicated, the error patterns are certainly in no respect random. In general this channel has long, high-density bursts followed by much longer error-free intervals.

The data implies that consideration should be given to error correction by retransmission and hybrid means. The presence of long bursts tends to preclude use of short block coding unless long delay interleavers are used.

DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)* The MITRE Corporation Bedford, Massachusetts	2a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED 2b. GROUP

3. REPORT TITLE

ERROR STATISTICS OF A MULTIPLE LINK DATA TRANSMISSION CIRCUIT

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*

N/A

5. AUTHOR(S) *(Last name, first name, initial)*

Brayer, Kenneth

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
February 1967	22	0

8a. CONTRACT OR GRANT NO. AF 19(628-5165 b. PROJECT NO. 705B c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) ESD-TR-66-337 9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)* MTP-40

10. AVAILABILITY/LIMITATION NOTICES

Distribution of this Document is Unlimited

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY Development Engineering Division Directorate of Aerospace Instrumentation L. G. Hanscom Field, Bedford, Mass.

13. ABSTRACT

Error patterns have been measured on a combined troposcatter, microwave, wireline circuit of more than 6000 circuit miles. The error data consist of more than 6×10^8 transmitted bits at rates of 1200 and 2400 bits/sec. These error patterns are presented in a form suitable for the analysis of error control, modem improvements or channel modeling.

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
SYSTEMS AND MECHANISMS Data Transmission Systems Multichannel Radio Systems Voice Communication (Troposcatter, Microwave, Wireline) Systems INFORMATION THEORY Coding MATHEMATICS Statistical Analysis (Troposcatter, Microwave, Wireline Error Locations) Statistical Distributions (Troposcatter, Microwave, Wireline Error Locations) Statistical Data (Troposcatter, Microwave, Wireline Error Locations)						

INSTRUCTIONS

1. ORIGINATING ACTIVITY: Enter the name and address of the contractor, subcontractor, grantee, Department of Defense activity or other organization *(corporate author)* issuing the report.

2a. REPORT SECURITY CLASSIFICATION: Enter the overall security classification of the report. Indicate whether "Restricted Data" is included. Marking is to be in accordance with appropriate security regulations.

2b. GROUP: Automatic downgrading is specified in DoD Directive 5200.10 and Armed Forces Industrial Manual. Enter the group number. Also, when applicable, show that optional markings have been used for Group 3 and Group 4 as authorized.

3. REPORT TITLE: Enter the complete report title in all capital letters. Titles in all cases should be unclassified. If a meaningful title cannot be selected without classification, show title classification in all capitals in parenthesis immediately following the title.

4. DESCRIPTIVE NOTES: If appropriate, enter the type of report, e.g., interim, progress, summary, annual, or final. Give the inclusive dates when a specific reporting period is covered.

5. AUTHOR(S): Enter the name(s) of author(s) as shown on or in the report. Enter last name, first name, middle initial. If military, show rank and branch of service. The name of the principal author is an absolute minimum requirement.

6. REPORT DATE: Enter the date of the report as day, month, year; or month, year. If more than one date appears on the report, use date of publication.

7a. TOTAL NUMBER OF PAGES: The total page count should follow normal pagination procedures, i.e., enter the number of pages containing information.

7b. NUMBER OF REFERENCES: Enter the total number of references cited in the report.

8a. CONTRACT OR GRANT NUMBER: If appropriate, enter the applicable number of the contract or grant under which the report was written.

8b, 8c, & 8d. PROJECT NUMBER: Enter the appropriate military department identification, such as project number, subproject number, system numbers, task number, etc.

9a. ORIGINATOR'S REPORT NUMBER(S): Enter the official report number by which the document will be identified and controlled by the originating activity. This number must be unique to this report.

9b. OTHER REPORT NUMBER(S): If the report has been assigned any other report numbers *(either by the originator or by the sponsor)*, also enter this number(s).

10. AVAILABILITY/LIMITATION NOTICES: Enter any limitations on further dissemination of the report, other than those imposed by security classification, using standard statements such as:

(1) "Qualified requesters may obtain copies of this report from DDC."

(2) "Foreign announcement and dissemination of this report by DDC is not authorized."

(3) "U. S. Government agencies may obtain copies of this report directly from DDC. Other qualified DDC users shall request through
_______________________________."

(4) "U. S. military agencies may obtain copies of this report directly from DDC. Other qualified users shall request through
_______________________________."

(5) "All distribution of this report is controlled. Qualified DDC users shall request through
_______________________________."

If the report has been furnished to the Office of Technical Services, Department of Commerce, for sale to the public, indicate this fact and enter the price, if known.

11. SUPPLEMENTARY NOTES: Use for additional explanatory notes.

12. SPONSORING MILITARY ACTIVITY: Enter the name of the departmental project office or laboratory sponsoring *(paying for)* the research and development. Include address.

13. ABSTRACT: Enter an abstract giving a brief and factual summary of the document indicative of the report, even though it may also appear elsewhere in the body of the technical report. If additional space is required, a continuation sheet shall be attached.

It is highly desirable that the abstract of classified reports be unclassified. Each paragraph of the abstract shall end with an indication of the military security classification of the information in the paragraph, represented as *(TS), (S), (C), or (U)*.

There is no limitation on the length of the abstract. However, the suggested length is from 150 to 225 words.

14. KEY WORDS: Key words are technically meaningful terms or short phrases that characterize a report and may be used as index entries for cataloging the report. Key words must be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location, may be used as key words but will be followed by an indication of technical context. The assignment of links, rules, and weights is optional