PROJECT TECHNICAL REPORT

TASK 707

SPACE COMMUNICATION SYSTEM SPECIFICATION AND

PERFORMANCE CRITERIA STUDY

NAS 9-8166

2. September 1971

Prepared for

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
MANNED SPACECRAFT CENTER
HOUSTON, TEXAS

N71-35309
(ACCESSION NUMBER)

104
(PAGES)

CR-115156
(NASA CR OR TMX OR AD NUMBER)

FACILITY FORM 602

(THRU)

63
(CODE)

07
(CATEGORY)

Prepared by

Communications and Sensor Systems Department
Electronics Systems Laboratory

TRW.

SYSTEMS GROUP

17618-H203-R0-00

PROJECT TECHNICAL REPORT
TASK 707

SPACE COMMUNICATION SYSTEM SPECIFICATION AND
PERFORMANCE CRITERIA STUDY

NAS 9-8166   2 September 1971

Prepared for
NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
MANNED SPACECRAFT CENTER
HOUSTON, TEXAS

Prepared by
J. C. Miller

Approved by: _W̳B̳W̳a̳r̳r̳e̳n̳_
W. B. Warren, Task Manager

Approved by _____
John DeVillier, Manager
Communications and Sensor Systems
Department

TRW
SYSTEMS GROUP

CONTENTS

CONTENTS (Continued)

iv

CONTENTS (Continued)

TABLES

ILLUSTRATIONS

PRECEDING PAGE BLANK NOT FILMED

ILLUSTRATIONS (Continued)

x

ILLUSTRATIONS (Continued)

1   INTRODUCTION

This report summarizes the results of a study performed under Subtask II of Task 707 entitled "Space Communication System Specification and Performance Criteria Study "

The general objective of this subtask is to supply part of the solution of the problem facing NASA MSC in connection with the development of a new communication system for advanced spacecraft   The problem is that of ensuring that any new communication system meets expected performance criteria with a minimum of compatibility conflicts between subsystems, between spacecraft and ground terminals, and possibly between spacecraft and relay satellites   It is felt that if adequate performance criteria, together with standardized test configurations and standard test and calibration material can be provided to subsystem manufacturers and assembly and test facilities early in the development cycle, that expensive and time consuming changes and retrofits which have sometimes occured in past programs can be prevented

Figure 1 illustrates the features of a general test plan suggested as a guide for implementing a system of test standardization which could help realize this goal   The Manned Spacecraft Center and each off-site operation (manufacturers test areas, launch sites, etc) would be provided with a standard test configuration for each system tested   Magnetic recorder/ reproducers would be used to provide test and calibration signals, and to record test results on magnetic tapes or discs   For each system, a master tape or disc would be prepared containing identification, calibration signals and test signals   In addition to the data tracks, a timing signal track would be provided for precise speed control of the tape on playback   From the master copy, a file copy will be retained for use at MSC, and working copies will be sent to the off-site locations   For each system, the calibration signals would include test signals which have varying amounts of noise and distortion added to produce results which cover the expected range of conditions to be encountered in the systems to be tested   Thus,

MASTER TAPES
VOICE
ANNOTATION
CALIBRATION
SIGNALS
TEST SIGNALS

WORKING COPY

STANDARDIZED TEST EQUIPMENT

TEST SIGNALS

SYSTEM UNDER TEST

CALIBRATION RESULTS

TEST RESULTS

FILE COPY

TAPE RECORDER

FILE COPY

SYSTEM SIMULATION AND TEST EQUIPMENT

COPY

COPY

ASSEMBLY AND TEST AREA

COMPARE RESULTS

RELEASE SYSTEM UNDER TEST

STANDARDIZED TEST EQUIPMENT

SYSTEM UNDER TEST

TO ASSY AND TEST

CALIBRATION RESULTS

TEST RESULTS

FOR FLIGHT

COPY

TAPE RECORDER

FILE COPY

Figure 1   Test Flow Plan

the calibration signals would simulate the output of the system under test under various conditions with the standard test inputs applied. When these calibration signals are applied to the standardized test equipment at the field locations, the results should compare closely with the results obtained when the calibration signals are applied to the standardized test equipment at MSC

When the field locations have verified that their standardized test equipment results are in agreement with those at MSC, the test signals would be applied to the system under test and the results would be recorded on the tape or disc containing the results of the calibration, as well as being displayed at the test site. The tape or disc containing the calibration and test results would be returned to MSC where verification that the system under test has met the criteria can be made. In addition, the tape or disc provides a permanent record of the test which could be filed at MSC. When MSC had been satisfied that the system under test has met the criteria, the system could be released for shipment to the next test location, where the procedure would be repeated. At the flight test location, the procedure could be utilized for pre-flight readiness checks. Details of the calibration and test signals are provided in the sections on voice, video, and digital data

The more specific objectives of this subtask can be listed in the following general categories

1. Investigation of the adequacy of criteria presently used or proposed to specify and evaluate communication systems in the areas of voice, video, and data transmission

2   Development of new performance and test criteria in those areas where current criteria are not considered adequate.

3   Development of representative methods of mechanizing tests to determine compliance with specific selected criteria.

After consideration of criteria presently used (Reference 1) a number of methods of specifying criteria were examined in more detail and were compared on the basis of their performance in the following five categories (Reference 2)

1. Precision of test data
2. Conciseness of results
3. Required data reduction
4. Pertinent to actual system performance
5. Ease of simulation

As a result of this comparison, three methods of specifying criteria were selected for more detailed study - one each for voice, video, and digital data systems. Suggested methods of mechanizing the implementation of these methods are included in Section 4

2. BACKGROUND SURVEY OF PERFORMANCE CRITERIA

2.1 GENERAL DISCUSSION

2 1.1 Introduction

A common problem in the establishment of performance criteria for communication transmission systems is the availability of widely accepted standards against which candidate criteria can be evaluated. The standards must incorporate sensitivity to all types of transmission impairments and in the case of speech or picture information, must relate these impairments to their subjective effects on the human user. This section summarizes the results of a survey of performance criteria conducted in the areas of voice, television picture quality, and digital data.

2 1.2 Voice Criteria

Speech transmission and processing systems have been the subject of intensive study over the past forty years and a large quantity of experimental data is available on the performance of such systems under a wide range of transmission and processing impairments Over the years, a general method of articulation testing has evolved to serve as a means of evaluating the subjective effects of these transmission impairments More recently, the detailed procedures for performing articulation testing have, to a large extent, become standardized and are the subject of a United States Standards Association document The detailed procedures contained in this document constitute an acceptable standard against which non-subjective performance criteria for speech transmission systems can be evaluated A brief review of the subject of articulation testing and its relationship to word intelligibility and signal-to-noise ratio is contained in Section 2.2. It is concluded that the use of the articulation index as a means of determining speech intelligibility should be considered as a prime candidate for speech criteria

2 1 3 Television Picture Quality Criteria

The investigations of the subjective effects of transmission impairments on picture perception are of more recent origin and not nearly so extensive

as those for speech transmission systems. Although most investigators have modeled their picture impairment investigations to a considerable degree on the previous work in speech systems, as yet there seems to be no wide agreement on the detailed testing arrangement that could be used to establish a standard of performance for picture transmission.

However, several agencies, including NASA MSC, the International Radio Consultative Committee (CCIR) and the Electronic Industries Association (EIA), do include a minimum required SNR as part of their television standards. These standards set limits on such parameters as the peak-to-peak picture signal level, polarity of the picture signal, frequency response of the video channel, as well as minimum requirements for picture signal-to-noise ratios (Reference 3) These SNR standards are generally set higher than the value at which noise would become just perceptible, and do not attempt to define a relationship between picture quality and SNR.

Since the question of a suitable standard for evaluating picture transmission impairments is still essentially unresolved, the study effort which forms the basis for this section has been largely concentrated in the area of picture evaluation. The purpose of Section 2.3 is to examine some of the methods which have been used to define picture quality of television signals Methods used by a number of researchers concerning the rating of subjective picture quality, viewing conditions, use of weighted noise, and the relationship of SNR to picture quality are discussed It is concluded that it is feasible to establish television quality criteria related to picture SNR

2 1 4 Digital Data Criteria

A brief discussion of the use of average bit error rate (BER) as a criteria for digital data systems is contained in Section 2.4, where techniques for shortening the measurement time for average BER are referenced. The problem involved in using average BER for systems with error correcting codes is also discussed, as well as a similar problem involved in acquiring synchronizing information.

2 2 VOICE INTELLIGIBILITY CRITERIA

2.2 1 Articulation Testing

In attempting to set criteria for voice communication an obvious choice
is a criterion which measures the ability of an individual receiving the
communication to understand it, i.e., how well can the individual receiving
the communication repeat the message content transmitted? One method of
implementing such a criterion makes use of articulation testing to establish
how well a communication is understood

Articulation testing is often used as a means of evaluating the per-
formance of speech transmission systems   A typical testing arrangement
includes the test material in the form of a prerecorded tape which is fed
into the system under test   The output of the system under test is recorded
and then played back into an audio system where it is listened to by a panel
of trained listeners   The method used and typical word lists are contained
in "Monosyllabic Word Intelligibility," a standard produced by the U S A
Standards Institute (Reference 4).

The listeners record the test material they hear   Their record is
compared with the transmitted material and the percent correct received
material is taken as a measure of voice intelligibility transmitted through
the system under test. In order to obtain uniform results from this test
procedure, standardized test material must be used. Although sentences and
nonsense syllables have been used at times, by far the most widely used
material is isolated words which are phonetically balanced

Some ground rules that are commonly employed in performing articulation
tests are

1. A fixed length list of test words is used   One length commonly
    employed is 50 words long and is referred to as the "PB 50 word
    list."

2   The listening team has practiced thoroughly with the list to be
    used in the test and have memorized the words in the list

3. Listeners always make a forced choice decision of which one of the test words was sent

4. Only percent words correctly understood, i.e., intelligibility, is of interest in determining the final score   No weight is given to lack of speaker identification or emotional content of the received test material

2 2.2 Articulation Index

Because of the difficulty in performing actual articulation tests with listeners, it is often desireable to employ some more easily measured quantity which is monotonically related to intelligibility   One such per-formance indicator is signal-to-noise ratio. The computational link between signal-to-noise ratio and word intelligibility is the Articulation Index, commonly abbreviated AI.

The concept of AI is based on the experimentally observed fact that speech signals can be divided into frequency bands in such a way that the speech information in each band has an equal effect on the intelligibility of the overall speech signal. The deletion of any one of the bands will degrade the intelligibility of the speech signal in the same manner as would the deletion of any other band. It has been found that 20 bands is a sufficient number of bands to describe the range from 200 to 6100 Hz   These bands are listed in Table 1   It has also been established that numbers representing the articulation values for each band can be considered as probabilities (Reference 5). Thus, the overall articulation can be repre-sented as the product of the individual articulation values of each band This fact allows AI to be defined as the sum of weighting functions (w_1) which are related to the signal-to-noise ratios of each band, since they are expressed as decibels

The weighting functions w_1 applied to each equal importance frequency band depend on the observation that if the SNR in a particular band is less than -12 dB, the effect of that band on the overall articulation index (AI)

Table 1

Equal Importance Bands for Normal Speech

Articulation Band Number	Frequency Range in Hz	Bandwidth in Hz
1	200 - 330	130
2	330 - 430	100
3	430 - 560	130
4	560 - 700	140
5	700 - 840	140
6	840 - 1000	160
7	1000 - 1150	150
8	1150 - 1310	160
9	1310 - 1480	170
10	1480 - 1660	180
11	1660 - 1830	170
12	1830 - 2020	190
13	2020 - 2240	220
14	2240 - 2500	260
15	2500 - 2820	320
16	2820 - 3200	380
17	3200 - 3650	450
18	3650 - 4250	600
19	4250 - 5050	800
20	5050 - 6100	1050

can be disregarded, and also that the effect of increasing the SNR in a particular band beyond +18 dB has no additional affect For each band the values of SNR between -12 dB and +18 dB have a linear effect on AI, resulting in the values of w_i given below.

$$w_i = 0 \quad \text{if } (S/N)_i < -12 \text{ dB}$$

$$w_i = \frac{(S/N)_i + 12}{30} \quad \text{if } -12 \text{ dB} \le (S/N)_i \le 18 \text{ dB} \tag{1}$$

$$w_i = 1 \quad \text{if } (S/N)_i > 18 \text{ dB}$$

where $(S/N)_i$ is the SNR in the ith band expressed in dB

The articulation index (AI) is then computed from

$$AI = \frac{1}{N} \sum_{i=1}^{N} w_i \tag{2}$$

where N = number of articulation bands in the pass band of the system under test $N = 20$ for the pass band of normal speech ranging from 200 to 6,100 Hz

For a system whose pass band is restricted to something much less than 6,100 Hz, the values of some of the w_i will be zero Since the remaining w_i are always in the range $0 \le w_i \le 1$, the AI for such a system will always be less than 1 even for very large SNR The corresponding percent word intelligibility will always be less than 100%

2.2 3 Relationship of Articulation Index and Voice Intelligibility

Unfortunately there is no direct computational procedure linking AI and word intelligibility, and the relationship has been established empirically This empirical relation was established be means of listening tests performed under various signal-to-noise conditions The intelligibility values were obtained by noting the percent of the transmitted test material that was correctly received by the listeners The test materials used

included words, sentences, and syllables. The calibration curves obtained for the different types of test materials are shown in Figure 2 One important thing to notice about these curves is that for a given AI the corresponding value of intelligibility is a strong function of the test material employed The correct curve for use with any particular system would depend on the length of the word list employed in the test. Until the length of the list used in specifying the word intelligibility requirements for the system can be determined, no corresponding Articulation Index requirements can be established

Once a particular calibration curve has been specified, it is possible to use this curve to compute word intelligibility for the communication system in terms of the power spectra of the signal and noise appearing at the output of the system The procedure for performing this calculation is as follows

1. Determine the signal-to-noise ratio in each of the 20 equally important bands covering the range 200 to 6,100 Hz If the bandpass of the system under consideration is restricted to something less than 6,100 Hz, then the signal-to-noise ratio in those bands not covered by the system response is set equal to zero, resulting in $w_i = 0$ for those bands.

2. Use the values of signal-to-noise ratio in each band to compute the values of the w_i by means of Equation 1

3 Substitute these values of w_i in Equation 2 to compute AI

4 Using the appropriate calibration curve in Figure 2, find the value of word intelligibility corresponding to this given value of AI

2.2 4 Relationship of Articulation Index, Word Intelligibility and SNR

Since the AI, and hence the percent word intelligibility (PWI), depends on the spectral distribution of the signal and noise over the 20 articulation bands, it is apparent that two voice signals having the same overall signal-to-noise ratio might have different PWI. Some feeling for the magnitude
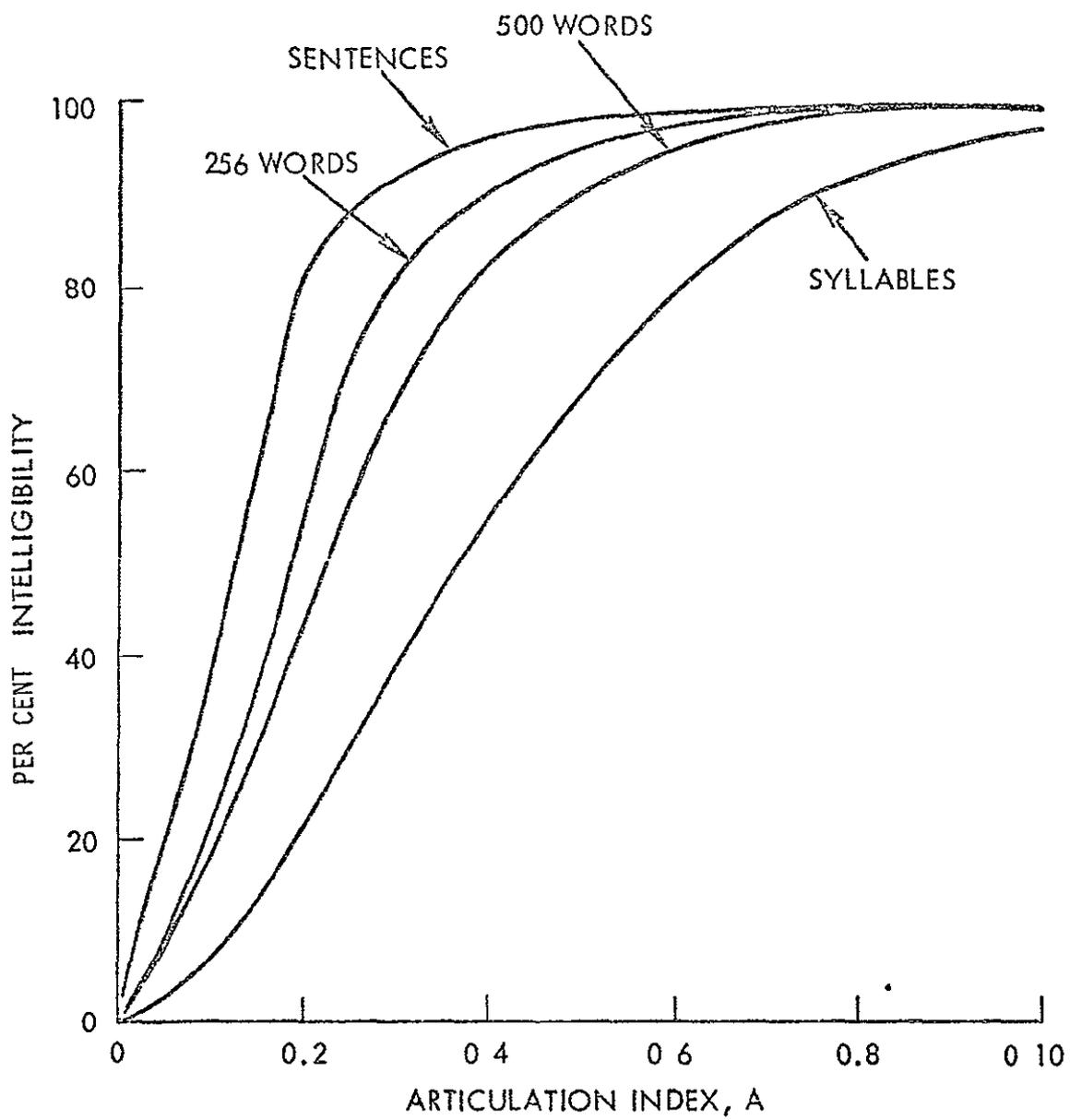
Figure 2   Intelligibility Versus Articulation Index

of this effect can be obtained from the results shown in Figure 3   These curves show the computed PWI versus SNR for both normal speech and speech equalized to have a flat power spectrum, both in the presence of white Gaussian noise   The pass band of the system assumed in making these calculations was 0-3 kHz, so that only the first 15 articulation bands were used in the computation   The results show that at least for these two particular speech spectrums, the PWI versus SNR relationship for normal voice spectral distribution is similar to that produced by "flat" speech

On the other hand, use of articulation index in computing PWI shows that PWI is sensitive to relative spectral distribution of speech and noise in the pass band   Curve B of Figure 4 shows that noise concentrated in higher end of the spectrum has little effect on AI and hence PWI   Curve B results from computing the AI assuming normal voice frequency distribution and use of 500 word list over a 0-3200 Hz band, in the presence of hypothetical noise that is flat from 2000 Hz to 3200 Hz and has very steep rise ($\approx$ 35 dB per octave)   This noise band is illustrated in Figure 5 together with a graph of normal voice spectrum   With an average SNR of -20 dB the percent intelligibility is still approximately 76% in the presence of the high frequency noise, thus illustrating the point that use of the AI gives a better indication of voice intelligibility that might be assumed from knowledge of the SNR alone

2 2 5 Summary and Conclusions

Results of the literature survey conducted indicate the desirability of speech criteria based on the determination of articulation index

(1) Articulation testing has been the subject of interest, research and testing over a long period of time   The method of determining AI and percent word intelligibility as described in Section 2 2 2, thus is one in which considerable confidence may be placed

(2) The method of determination of articulation index considers the spectral distribution of both the voice signal and the noise throughout the pass band   This method reduces erroneous conclusions which may result if an average signal to noise ratio is
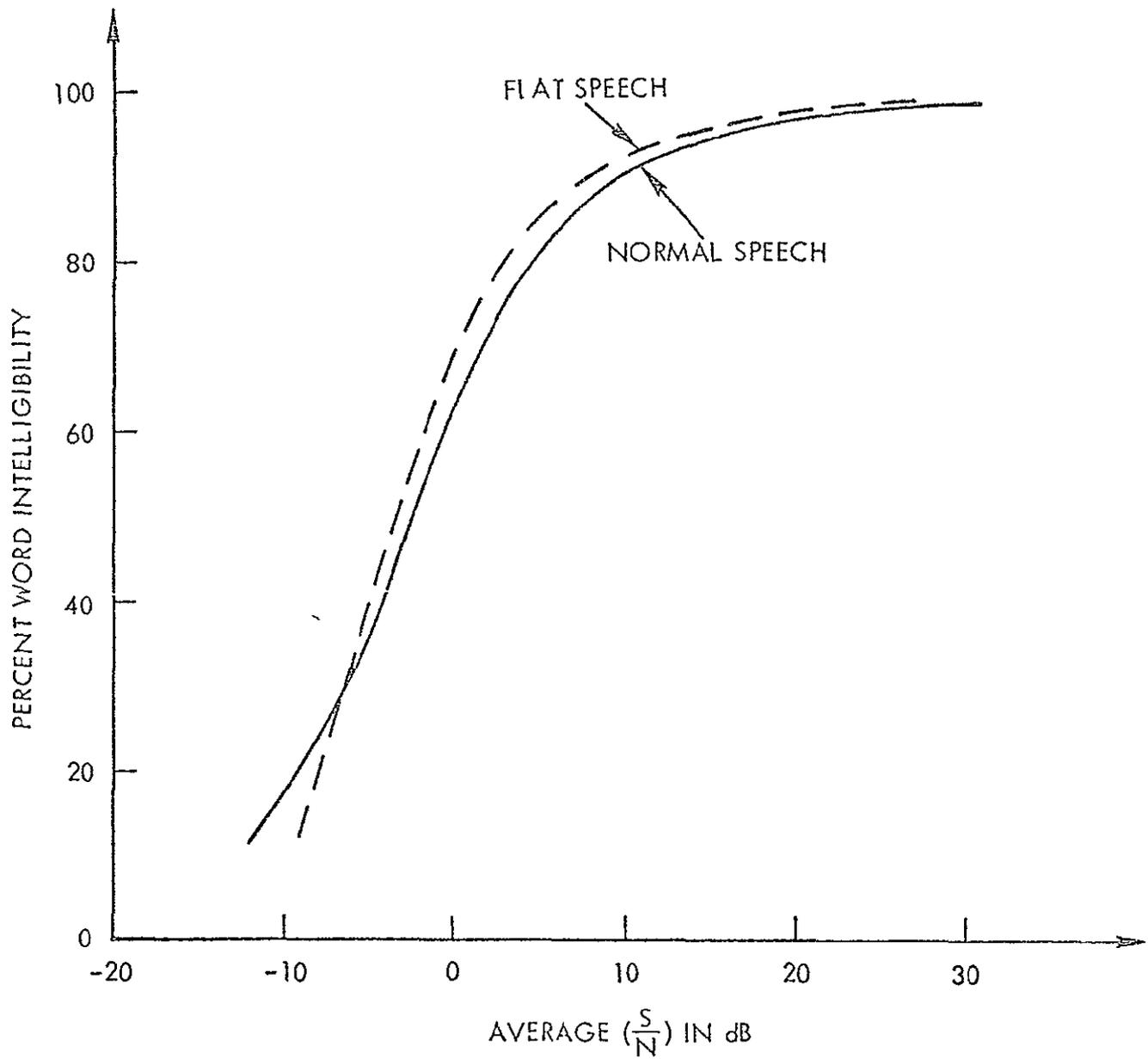
Figure 3. Intelligibility for Two Different Speech
Power Spectrums
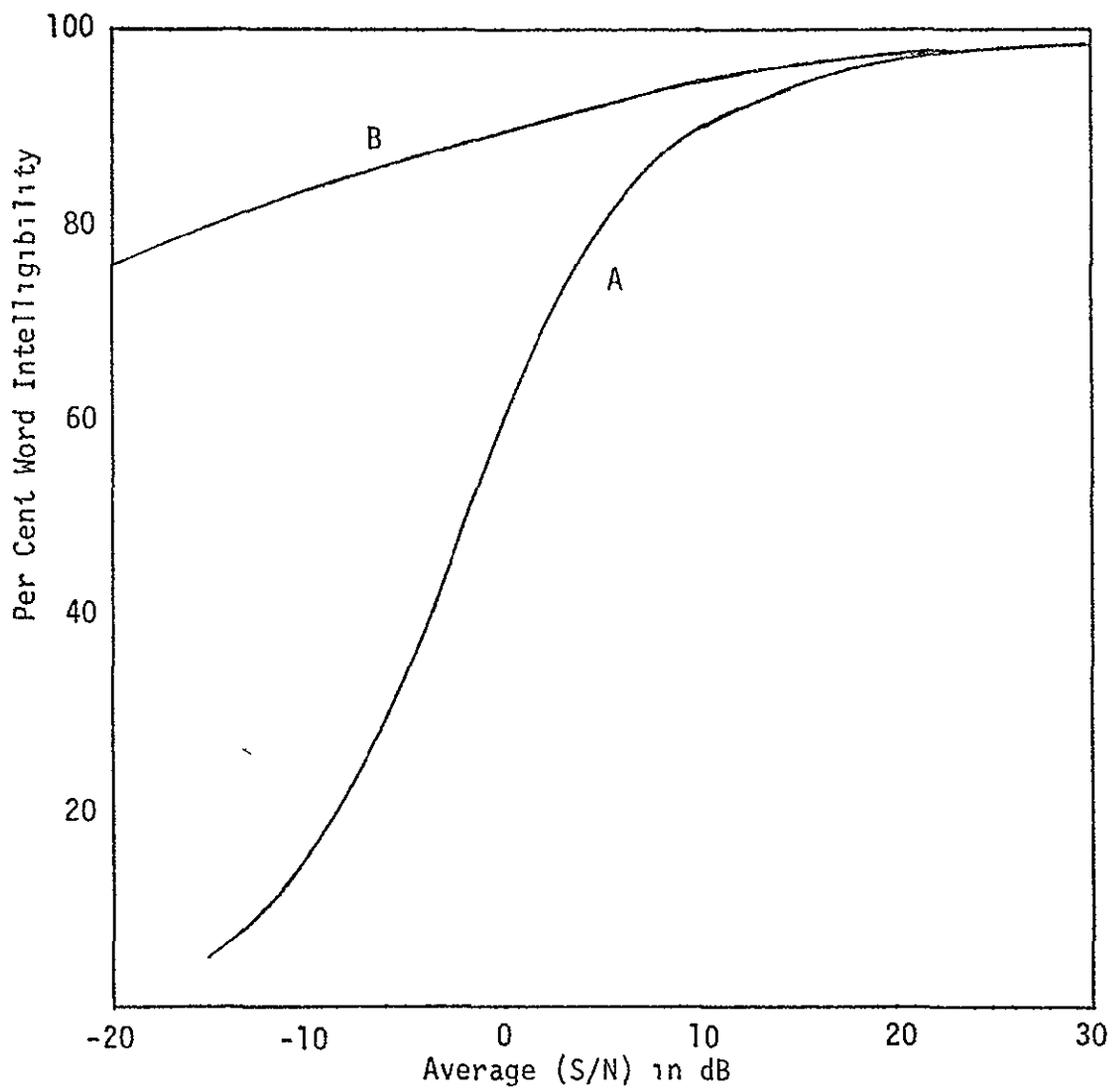
Normal Speech

Effect of Noise Spectral Distribution

Curve A   Flat Noise

Curve B   Noise Concentrated at High Frequency
End of Voice Band

Figure 4
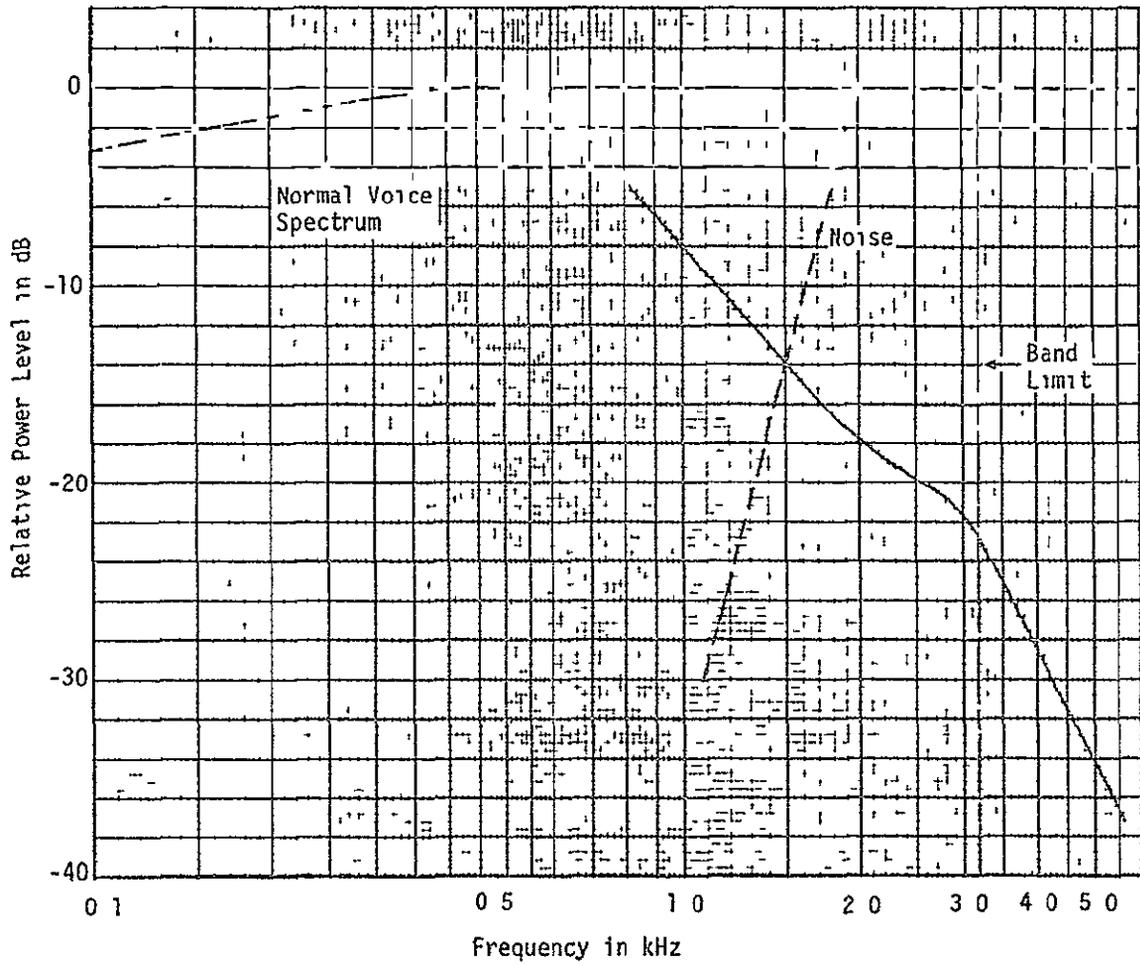
Figure 5   Normal Voice and High Frequency Noise

determined for the voice band and the voice signal and noise power happen to be concentrated at opposite ends of the pass band

The use of articulation index as a standard for voice criteria does present some problem areas

(1) Implementation of test methods involving articulation index may be a problem, especially if automation or semi-automation is desired   Speech to noise ratios for fifteen or twenty frequency bands in the voice pass band must be determined

(2) The method does not take into account the time relationship of the signal and the noise   Noise interspersed between speech sounds could give erroneously low articulation indices

2 3 CRITERIA FOR EVALUATING TELEVISION PICTURE QUALITY

2 3 1 Quality Rating Scales

The basic problem associated with any method of specifying criteria for television pictures based on subjective evaluation of picture quality is that of determining the quality rating or index to be associated with various amounts and kinds of impairments   Involved in this rating are the type of testing involved, class of observers used, test material (kinds of pictures) and viewing conditions   Only when these factors have been determined, can the relationship of the quality standards and some measurable physical quantity such as picture signal to noise ratio be determined   The purpose of this section is to examine some of the work that has been done in this area   As the authors of Reference 6 point out, the unfortunate fact is that the several investigations which have been made into this problem used different quality "indices", different classes of observers, and different viewing conditions

Three types of psychometric methods have been used in determining picture quality in television testing

1. Comparison - The magnitude of one kind of impairment is varied until it is judged to be equal to a fixed magnitude of another kind of impairment used as a reference

2-13

2. Discrimination - Which seeks to establish the threshold at which an impairment just becomes visible.

3 Opinion-rating - Different magnitudes of impairment are applied in random order to a picture and the observer rates them by selecting one value of a pre-determined scale

The creation of a rating "code" or "index" is one of the first problems requiring solution in psychrometric testing Most of the scales used in the past whether for "impairment" type testing or "quality" (opinion-rating) type testing involve the use of six or seven levels of quality. Investigators at the Bell Laboratories and at the British Broadcasting Corporation (BBC) have used impairment scales which exhibit substantial agreement·

Bell Laboratories "Impairment Scale"

Comment Number

1. Not perceptible

2. Just perceptible

3. Definitely perceptible, but only slight impairment to picture

4. Impairment to picture, but not objectionable

5. Somewhat objectionable

6. Definitely objectionable

7. Extremely objectionable

BBC "Impairment" Scale

Score

1 Imperceptible

2. Just perceptible

3. Definitely perceptible but not disturbing

4 Somewhat objectionable

5. Definitely objectionable

6. Unusable

It is to be noted that the two scales would be almost identical if comment number 4 were deleted from the Bell Laboratories Scale.

In addition, BBC has used a "quality" scale

Score

1. Excellent
2 Good
3 Fairly Good
4 Rather Poor
5. Poor
6. Very Poor

R D Prosser, et al (Reference 6) have proposed a simple quality scale of 5 "grades" with the same text as the BBC scale, but with grades 3 and 4 combined into one grade labeled "Fair"

A - Excellent
B = Good
C - Fair
D - Poor
E - Bad

A scale or index combining both "impairment" and "quality" was used by the Television Allocation Study Organization (TASO) in tests conducted in 1958

Number	Name	Description
1	Excellent	The picture is of extremely high quality, as good as you could desire.
2.	Fine	The picture is of high quality providing enjoyable viewing Interference is perceptible
3	Passable	The picture is of acceptable quality Interference is not objectionable
4	Marginal	The picture is poor in quality and you wish you could improve it. Interference is somewhat objectionable
5	Inferior	The picture is very poor but you could watch it Definitely objectionable interference is present
6	Unusable	The picture is so bad that you could not watch it

Since "still" pictures seem to inspire more critical judgement than do moving pictures, the test material used in most tests were colored slides (Reference 6, 7, 8).

2 3.2 Viewing Conditions

As regards viewing conditions, the ambient illumination in the test areas was quite low, generally the light from the picture tube reflected from light colored walls and ceilings The picture peak luminance ranged from 10 to 30 foot-lamberts in the highlights The parameter of observer viewing distance was usually expressed in terms of "viewing ratio" (distance divided by picture height), the ratio used varied from 4 (Reference 7) to 8 (some TASO tests)

2 3 3 Relationship of Picture Quality and SNR

Having determined a suitable test method (comparison), and a scale or index to be used with the tests, the next requirement is to determine the relationship between the subjective measurements and some objective standard Use of signal-to-noise ratio (SNR) as a standard of comparison is an obvious choice, and has been used in most of the studies considered

Use of SNR, however, raises several questions The SNR is to be measured in what portion of the television receiver circuit - RF input or video? What is the noise bandwidth? What is noise spectral distribution? How well does the test configuration simulate a real world television link?

Since picture quality impairment is directly related to the level and spectral distribution of noise in the video channel, it is common to use picture SNR as objective standard against which the subjective quality ratings are compared This picture SNR can be defined as

$$(S/N)_p = \left(\frac{\text{blank-white video voltage}}{\text{RMS voltage of video noise}}\right)^2 \qquad \text{(Reference 9)}$$

2-16

where the blank-to-white video voltage range is from 0 to 0 714 volts, and
the RMS noise bandwidth is the nominal video bandwidth of 4 2 MHz (in the
American Standard NTSC system)   In the NTSC system, each television chan-
nel occupies a 6 MHz bandwidth with the picture carrier (amplitude modulated)
separated from the sound carrier (frequency-modulated) by 4 5 MHz   In
color broadcasting a color subcarrier is spaced 3 58 MHz from the picture
carrier

The most common method of testing uses a flying spot scanning device
to create the desired video signal with the noise added in the video channel
to create a picture SNR   However, one important series of tests, that
performed by TASO, utilized a standard closed circuit television test con-
figuration when RF noise was added at the receiver input resulting in RF
carrier-to-noise ratio, which must be converted to an equivalent picture
SNR before comparisons can be made with other test results   An example of
the direct picture SNR method is shown in Figure 6 (from Reference 7), and
a simplified version of the TASO test configuration is shown in Figure 7

2 3 4 Use of Noise Weighting

The concept of picture SNR as standard for subjective picture quality
rating is not very meaningful unless the spectral distribution of the noise
is known, since noise near the upper end of the video band is considerably
less objectionable than noise at the low end of the video band   Barstow
and Christopher of Bell Laboratories (Reference 7) have conducted a series
of experiments to determine a method by which random interference to tele-
vision pictures could be measured so that equal measured magnitude of noise
(RMS) would produce equal picture impairment. These tests, initially per-
formed in 1953 for monochrome television and repeated in 1962 with the
addition of color TV tests, resulted in the concept of weighted noise,
defined as video noise measured with a weighting network which represents
the perception of noise of different frequency bands by an average viewer

The test configuration shown in Figure 6 was used to determine the
characteristic of the noise weighting filter by means of the comparison
technique mentioned earlier. The test observer was instructed to adjust

2-17
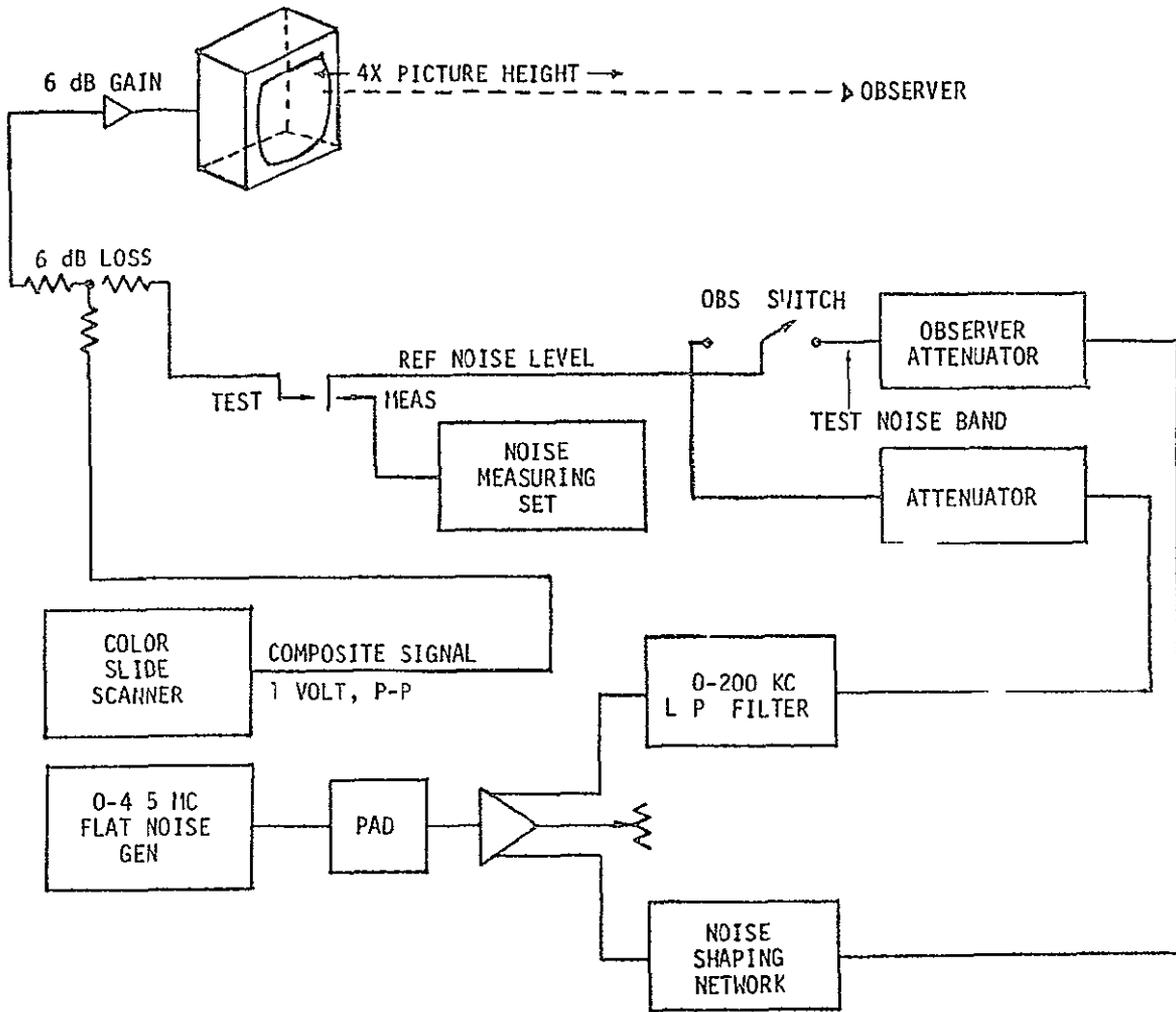
Figure 6. Bell Laboratories Equipment Arrangement for
Subjective Test of Video Interference (Reference 3)
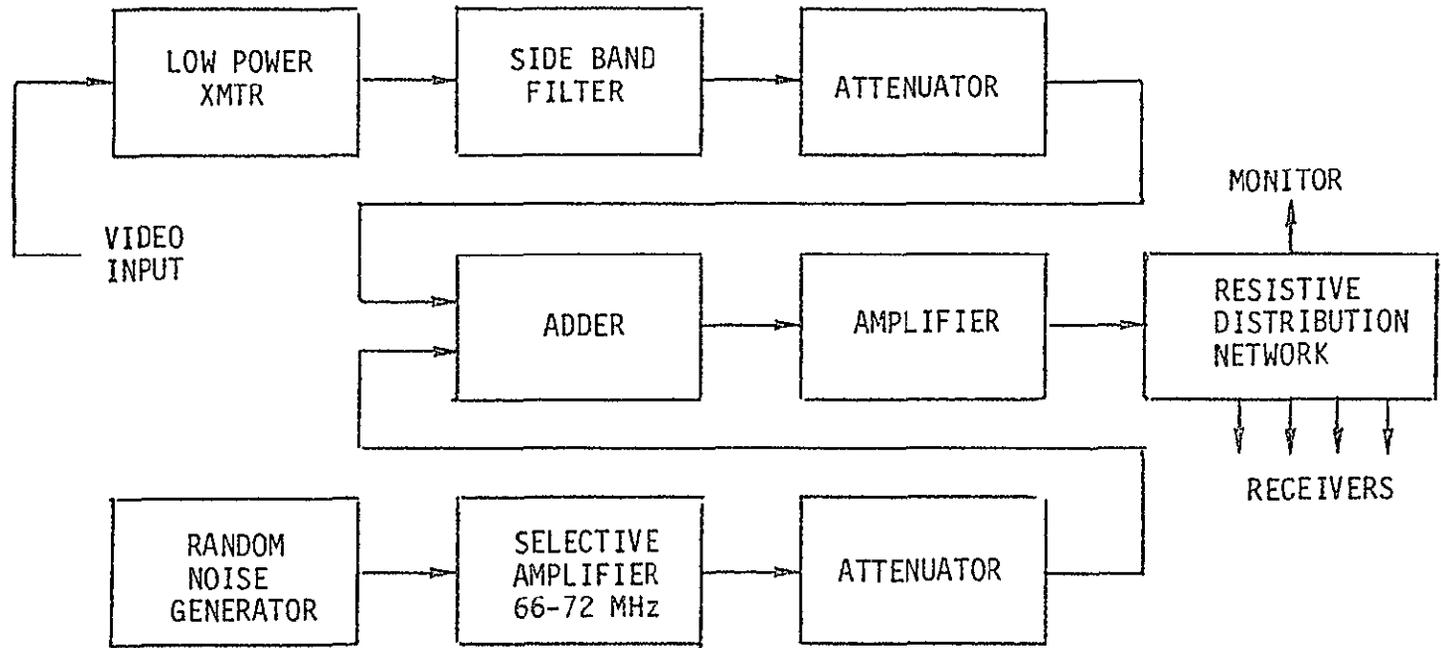
Figure 7. Simplified Diagram of TASO Random
Noise Tests (Reference 5)

the level of the test noise until the interfering effect was judged equal
to the standard (that passed through the 0-200 kHz L P filter). The test
results of a number of observers when averaged resulted in Figure 8 where
the weighting characteristics are normalized to 0 dB at zero frequency
The 1962 monochrome characteristic agrees very closely with the weighting
characteristic described by Jansen and Jorden in defining weighted picture
SNR (Reference 9) The "hump" in the color characteristic is due to inter-
ference with the color subcarrier at 3 58 MHz The test results used to
determine the weighting characteristic show that interfering effect of noise
near the color subcarrier is strongly dependent upon the subject picture,
for colors with low saturation approaching grey, or highly saturated bright
colors such as yellow, cyan or light blue, the "peak" would be approximately
5 dB lower than that shown, while for heavily saturated reds and blues the
peak would be 2 or 3 dB higher The "hump" shown is, in effect, a somewhat
conservative compromise between weightings of bright colors and heavily
saturated red and blue colors

As a test of the effectiveness of the weighted noise concept, Barstow
and Christopher conducted a series of companion tests using the test set up
of Figure 6 In the monochrome tests, a series of 3 or 4 pictures were
shown to a total of 30 observers (3 panels of 10 observers) for each of 12
different conditions of impairment due to 12 different bands of noise,
ranging from flat noise (0 to 4 5 MHz) to 500 kHz noise bands centered at
different frequencies within the television video spectrum The noise was
measured in each case by two RMS devices, one employing no weighting, and
one with the weighting shown in Figure 8 Compared to an SNR (peak-to-
peak signal to RMS noise) of 50 dB for the reference noise of 0-200 kHz,
the average SNR resulting from tests using the weighted noise was 49 4 dB
with a standard deviation 0 7 dB, while the average SNR for the unweighted
noise was 41.8 dB with a standard deviation of 4 3 dB, showing that use of
noise weighting results in substantially equal subjective effects correspon-
ding to equal magnitudes of noise with different spectral content For
the color tests, more different noise frequency bands were added, resulting
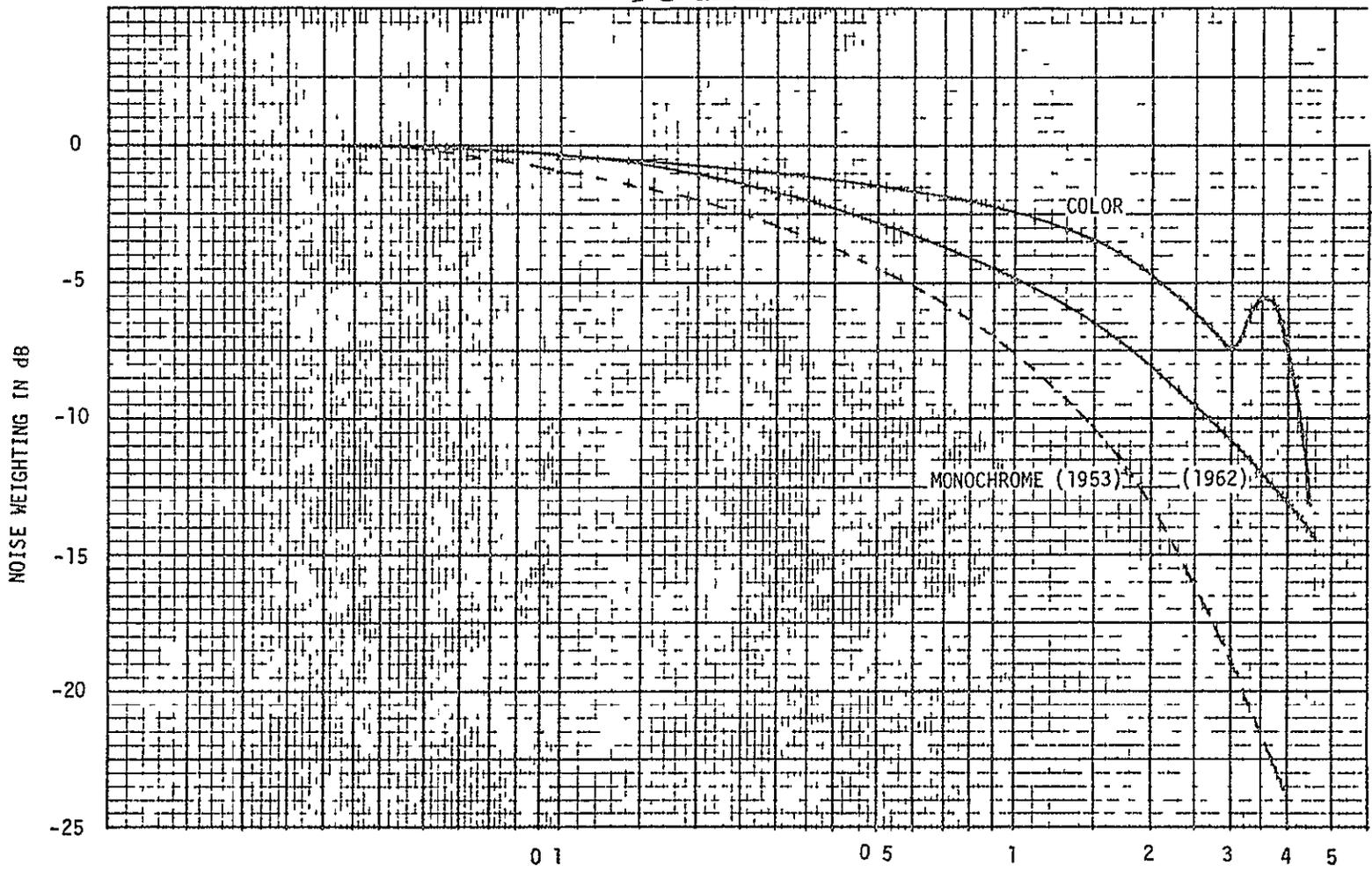
Figure 8   Video Frequency in MHz

in a total of 20 Compared to the 51.7 dB SNR of the reference noise, the average of the weighted data was 51.1 dB with a standard deviation of 0.7 dB, while the unweighted data average was 46 4 dB with a standard deviation of 3 2 dB Thus, it has been demonstrated that the use of weighted noise enables correlation of picture SNR with subjective picture quality criteria

2 3 5 Results of Quality Rating Tests Versus SNR

Utilizing the same test configuration, Barstow and Christopher ran a series of opinion rating tests using the Bell Laboratories "impairment" scale discussed earlier The results of two series of tests are shown in Figure 9, where the quality scores versus noise levels are plotted for the reference noise (0-200 kHz) and flat noise for both the monochrome and color tests The two curves are shown since they represent the extremes in noise bandwidth tested, and thus best illustrate the phenomenon that the slope of the curve of impairment resulting from narrow band noise as a function of noise amplitude is steeper than that for broad band noise This indicates that visual judgement of impairment does not follow exactly an RMS rule (Since the noise in each case was measured with an RMS device, the visual judgement of impairment for the narrow band noise versus noise magnitude should produce a curve parallel to that produced by flat noise) It is seen that this is very nearly the case in the monochrome tests and justifies the use of RMS devices in measuring noise in these types of tests. Even for the color tests, the effect when considered over the range of values of 2 "just perceptible" to 4 "not objectionable" (the range most generally of interest) the difference between the two curves increases from about 1 5 dB at the 2 level to approximately 3 dB at the 4 level, or about 1.5 dB difference while the noise magnitude increases approximately 8 dB from -53 dB to -45 dB

When attempts are made to compare the results of these "quality rating" tests with similar tests such as the TASO tests, several difficulties are encountered. In addition to the differences in rating scales pointed out earlier, the results of the Bell Laboratories tests are plotted in terms

7 Extremely
  Objectionable

6. Definitely
   Objectionable

5. Somewhat
   Objectionable

4. Impairment But Not
   Objectionable

3. Definitely Perceptible
   Only Slight Impairment

2. Just Perceptible

1 Not Perceptible

Monochrome

0 to 200 KC
Noise Average
Observer

Flat Noise
Average
Observer

Weighted Noise Magnitude
db Below 1 Volt RMS

7. Extremely
   Objectionable

6 Definitely
  Objectionable

5. Somewhat
   Objectionable

4 Impairment But Not
  Objectionable

3 Definitely Perceptible
  Only Slight Impairment

2. Just Perceptible

1. Not Perceptible

Color

0 to 200 KC
Noise Average
Observer

Flat Noise
Average
Observer

Weighted Noise Magnitude
db Below 1 Volt RMS
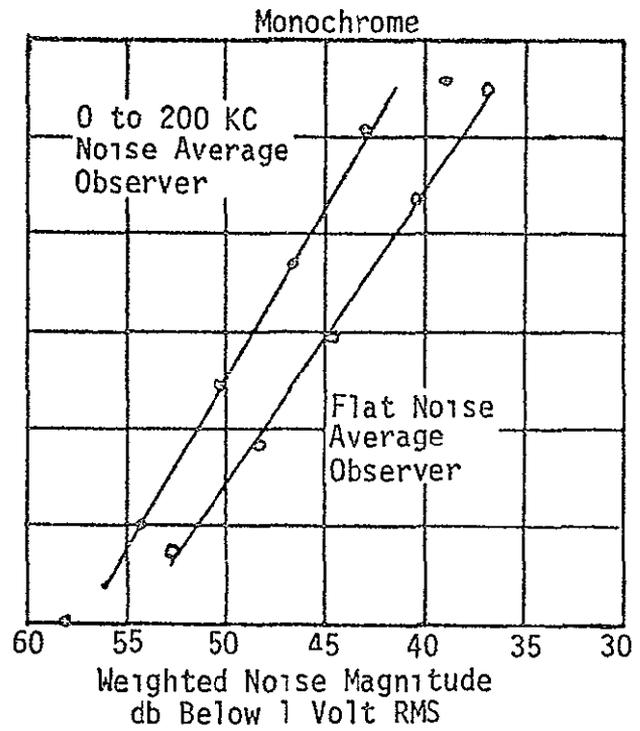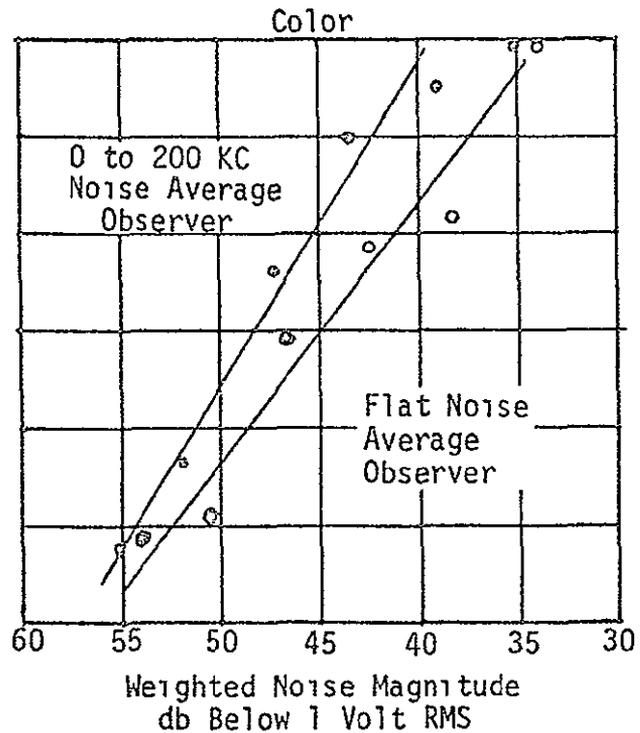
Figure 9. Bell Laboratories Quality Impairment
Test Results (Reference 3)

of noise magnitude (dB below 1 volt RMS) while the TASO results are given
in terms of SNR in dB. Also, the TASO results were orginally measured in
terms of RF carrier-to-noise ratios   J. Jansen, and P. L Jordan, in a
television broadcasting satellite study (Reference 9) manipulated the TASO
carrier-to-noise ratios to picture SNR values by including the effects of
an FM RF channel and TV camera noise. Their results are compared with the
Bell Laboratories results in Figure 10   The Bell Laboratories noise level
figures have been adjusted so that the weighted noise magnitude (dB below
1 volt RMS) corresponding to number one on the quality scale was made equal
to the picture SNR corresponding to the number one on the TASO quality scale
In addition, the six point TASO scale was adjusted so that each impairment
rating number was aligned with the rating in the Bell Laboratory scale to
which it corresponded most closely as regards word description   This
resulted in there being one point on the Bell Laboratories scale to which
there was no corresponding point on the TASO scale   Number 2 "perceptible
interference" on the TASO scale could be paired with either number 2 "just
perceptible" or number 3 "definitely perceptible" on the Bell Laboratories
scale   As shown on Figure 10, a straight line connecting points 1, 3, 4,
and 5 of the TASO scale passes approximately midway through points corre-
sponding to quality numbers 2 or 3 as plotted against the SNR value (40.3
dB) given as corresponding to TASO quality number 2 in Reference 2, sug-
gesting that 2 on the TASO scale has a subjective effect somewhere between
2 and 3 on the Bell Laboratory scale

Aside from the differences in the slope of the curve of the picture
quality versus SNR between the TASO and Bell Laboratories results, it is
interesting to compare the minimum or allowable SNR specified by different
groups and agencies which have set standards in the past   These minimum SNR
values are shown in Table 2   The NTSC and CCIR-I standards are taken from
Reference 9, as is the TASO number   The Bell Laboratories numbers are from
Reference 3 and actually are the noise magnitudes in dB below 1 volt RMS
that correspond to quality scale number one on the Bell Laboratories scale.
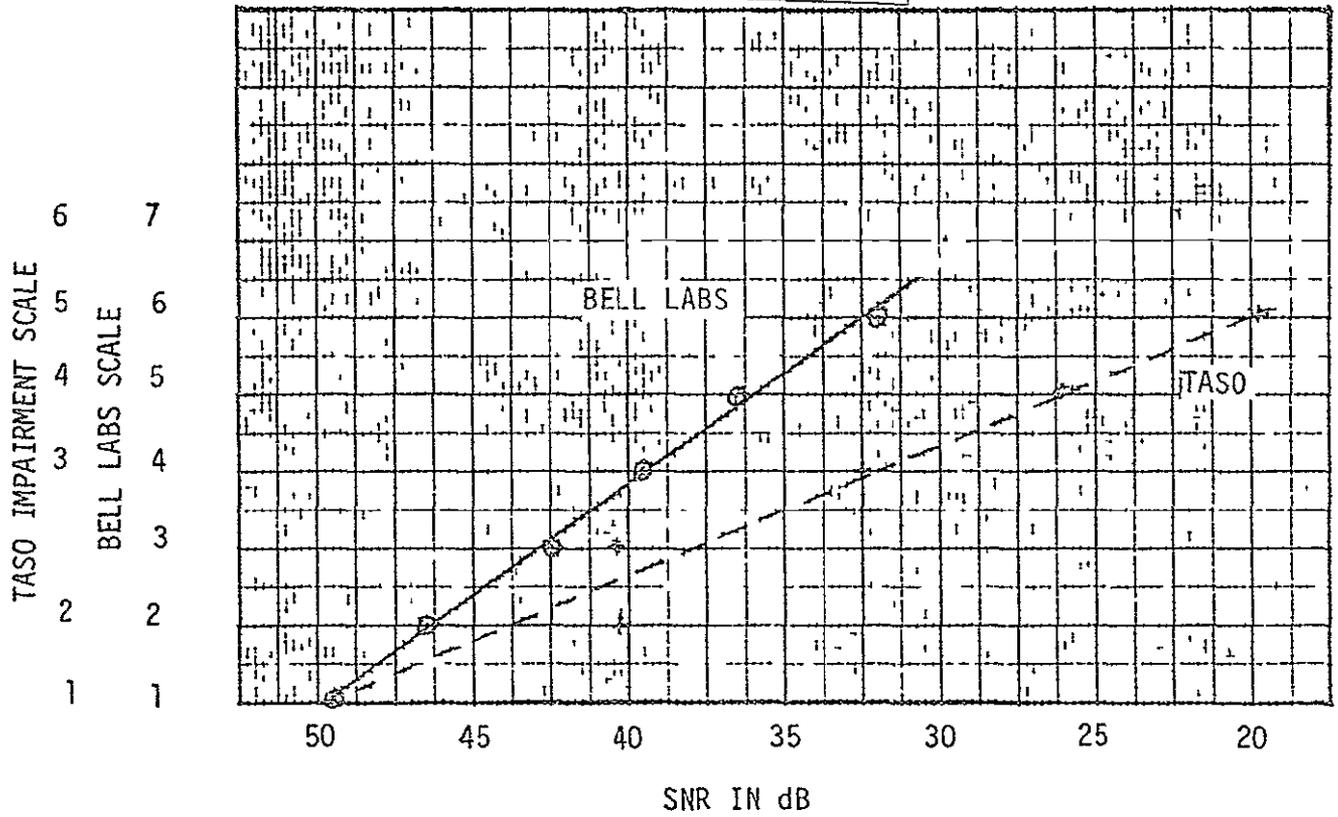
Quality Scales in Terms of Impairment

Bell Labs		TASO (Impairment Comments)	
1	Not Perceptible	1	Excellent
2	Just Perceptible Impairment	2	Perceptible Interference
3	Definitely Perceptible Slight Impairment		
4	Impairment but not Objectionable	3	Interference not Objectionable
5.	Somewhat Objectionable	4	Somewhat Objectionable Interference
6	Definitely Objectionable	5	Definitely Objectionable Interference
7	Extremely Objectionable	6	Unusable

Figure 10

The SNR values of the NTSC and CCIR standards, as well as the Bell Laboratories value are not based on consideration of source noise, such as television camera noise. The TASO value includes an allowance for this noise source.

Since the spread of values of SNR corresponding to the best quality or less perceptible impairment is about $\pm$ 4 5 dB from the average of the five values for monochrome TV, it would seem that the average of about 54 5 dB (picture SNR as defined on page 7) would be a reasonable value for good grade of television picture.

Table 2
Allowable Video (Picture) SNR

	NTSC 525 Line 4 2 MHz Bw	EIA	CCIR-I 625 Line 5.5 MHz	Bell Laboratories	TASO
Monochrome	56 dB	59 dB	52 dB	55 dB	49 5 dB
Color					
Composite Video	56 dB			57 dB	
Luminance Channel			52 dB		
Color Channel			46 dB		

2-26

2 3 6 Video Background Survey

Based on the literature survey conducted to date, several conclusions can be drawn

1   It appears feasible that a television picture quality rating scale can be incorporated in NASA specifications and test criteria.

2   Based on specified test conditions, a picture SNR can be defined, and a relationship between picture SNR and input RF signal level determined

3   A picture quality rating scale can be defined in terms of the picture SNR   Inherent in this relationship of picture quality versus SNR is the use of a noise weighting network which permits noise of equal magnitude to have the same visual subjective effect regardless of spectral composition

2 4 DIGITAL DATA CRITERIA

Techniques for evaluating the performance of digital data transmission systems are more widely agreed upon than those used to evaluate speech or picture transmission systems   The existence of this widespread agreement stems primarily from the fact that, for most data transmission systems, the human perception mechanism is not directly involved and subjective evaluations of the effects of transmission impairments are not required   The results of transmission impairments to digital data can be evaluated by performing a counting operation in which a decision is made concerning the correctness of incorrectness of each transmitted bit   By observing the resulting time sequence of errors over a sufficiently long interval, the error statistics of the transmission system can be determined   Although it is possible to measure all orders of error statistics by observation of the error sequence, in practice this process can be time consuming and difficult   The most readily obtained statistic is the average bit error rate which describes the probability of making N errors in a transmitted sequence of K bits   The attainment of a given average bit error rate is a

widely used criteria for specifying the performance of digital data transmission systems [10]. For systems where the distribution of errors in unimportant, the average bit error is an acceptable method of specifying system performance. The measurement technique commonly employed to measure average bit error is to let a large number of bits, say K, pass through the system under test and count the total number of errors, N, that occur The average bit error (BER) is then given by

$$BER = \frac{N}{K}$$

When the operating bit rate is low or the bit error rate is very small, an excessively long time may be required to obtain large enough values for N and K for the computed bit error rate to be statistically significant Techniques for shortening the measurement time for average bit error rate have been developed[11] These techniques make use of a multiple level decision mechanism to establish a pseudo-error rate. This pseudo-error rate can subsequently be used to estimate the true error rate Since the pseudo-error rate is larger than the true error rate, a shorter time is required to obtain a statistically significant measurement.

For some systems, average bit error rate is not an adequate criterion for specifying system performance because the error performance of these systems depends on both the number and distribution of errors in the transmission channel For example, a digital transmission system employing an error correcting code may be capable of correcting at the receiver any single or paired errors that occur in the transmission channel For such a system there is not necessarily any direct relationship which expresses the error rate the output of the system in terms of the average bit error in the channel Consequently, the specification of performance for such a system should not be made in terms of the average bit error rate in the transmission channel. Instead, the performance requirement should be imposed on the average bit error rate at the output of the receiver decoder. If

the decoder is not an integral part of the receiving equipment, the specification of the system performance can be written in terms of the higher order statistics of the data sequence appearing at the output of the systems decision device. For the example cited above, the statistic required is the probability of occurrence of triple and higher order errors. One method of obtaining this statistic is to arrange three counters which are driven by the error sequence. One counter totals the number of single errors, the second counter totals the number of paired errors, and the third counter totals all the errors regardless of type. If, in a long sequence of K bits, the first counter total is M, the second counter totals N, and the third counter totals P, the probability of making triple and higher order errors, P_e, is given by

$$P_e = \frac{P - (N + M)}{K}$$

Another situation in which the distribution of errors is important is in the acquisition and maintenance of synchronizing information. For example, an error which causes loss of frame synchronization is a more significant error than one which affects only one message bit. Unless the system is specifically designed to provide more protection against synchronization errors than message bit errors, the errors caused by loss of synchronization will be the determining factor in setting the overall error rate for the system. In such an event, average channel bit error rate is not as desireable a performance indicator as average frame synchronization error rate. If only frame synchronization errors are observed and it is desireable to specify the system performance in terms of average bit error rate, the technique discussed by Rohde[12] can be used to relate these two quantities. Rohde shows that the average bit error rate, P_e, is related to the frame synchronization error rate by

$$P_e = \sum_{i = 1}^{N} \frac{X_i}{KN}$$

where X_i is the number of bit errors in the ith sync word, N is the total number of words observed snd K is the number of bits in each sync word

It is concluded that while the average bit error rate (BER) is sufficient for situations where the distribution of errors is not important, the increased use of coding techniques in digital communications requires that criteria be written in terms of higher order statistics, depending upon the error correction capabilicies of the system under consideration For systems where the actual BER is so low that it causes difficulties in measuring, pseudo-error rates which are larger than the actual rates may be generated, from which the actual rates may be determined.

3. NEW CRITERIA DEVELOPMENT AND SELECTION

3.1 GENERAL DISCUSSION

The previous section reviewed performance criteria presently used in evaluating voice, video, and digital data systems   The conclusions reached were that measurement of articulation index for voice systems, weighted picture SNR for video systems, and measurement of BER for digital data systems offered the most promising areas for continued investigation Within these broad categories, different methods of specifying criteria are possible   In addition, several methods which do not fall within these categories are considered

The purpose of this section is to describe and evaluate a number of methods of specifying criteria   The method employed is to propose several candidate evaluation systems in each category, and make an evaluation of each system based on a number of parameters   precision of test data, conciseness of results, required data reduction, pertinence to actual system performance, and ease of simulation   Each of these parameters is assigned a relative weighting value.   In order to evaluate the candidate systems, the factors affecting each of the performance parameters of each of the candidate systems are considered, and a numerical rating value is assigned to each   By this procedure an overall score for each of the candidate systems is achieved which will provide the basis for recommendations concerning these methods of specifying performance criteria   Before these criteria are applied in the evaluation of a specific system, the numerical ratings assigned in the trade-off matrix should be re-examined to determine if more specific information is available which is sufficient to warrant the selection of a different test criteria

The discussion of each candidate system includes a description of the method involved in mechanizing the test and computing the criteria, and a simplified block diagram of the process   A least one method, Method B for television, represents an unproven principle - that video signals can

be separated into frequency bands having equal subjective importance to the viewer, in a manner similar to the principle involved in determining the articulation index of a voice system

Two of the candidate systems, Methods C and F for determining speech SNR for voice systems have been developed and tested under NASA MSC direction by the Philco-Ford Corporation (Reference 13 and 14).

3 2 VOICE SYSTEMS CRITERIA

3.2 1 Method A - Determination of Articulation Index

Method A described below is a method of mechanizing the testing to determine the AI of a voice communication system A block diagram is shown in Figure 11

Description of Method

A. The input tape is composed of a standard phonetically balanced word list scored for word intelligibility

B Modulation, RF source and RF link simulation are determined by the system under test If the test is an end-to-end test of a complete link, the modulator and RF source would be replaced with components of the system under test ·

C The input tape signal is fed into a variable equalizer to match the magnitude, $|A(\omega)|$, and phase, $\phi(\omega)$, of the system under test.

D The output signal S is fed through switch S_1 to a difference network, to which the output S+N from the system under test is also applied

E In switch position 1, only the output S+N from the system under test is fed to the difference network, resulting in an output S+N

F. In switch position 2, both the matched output S and the output S+N are fed to the difference network resulting in N, the system noise, as an output.
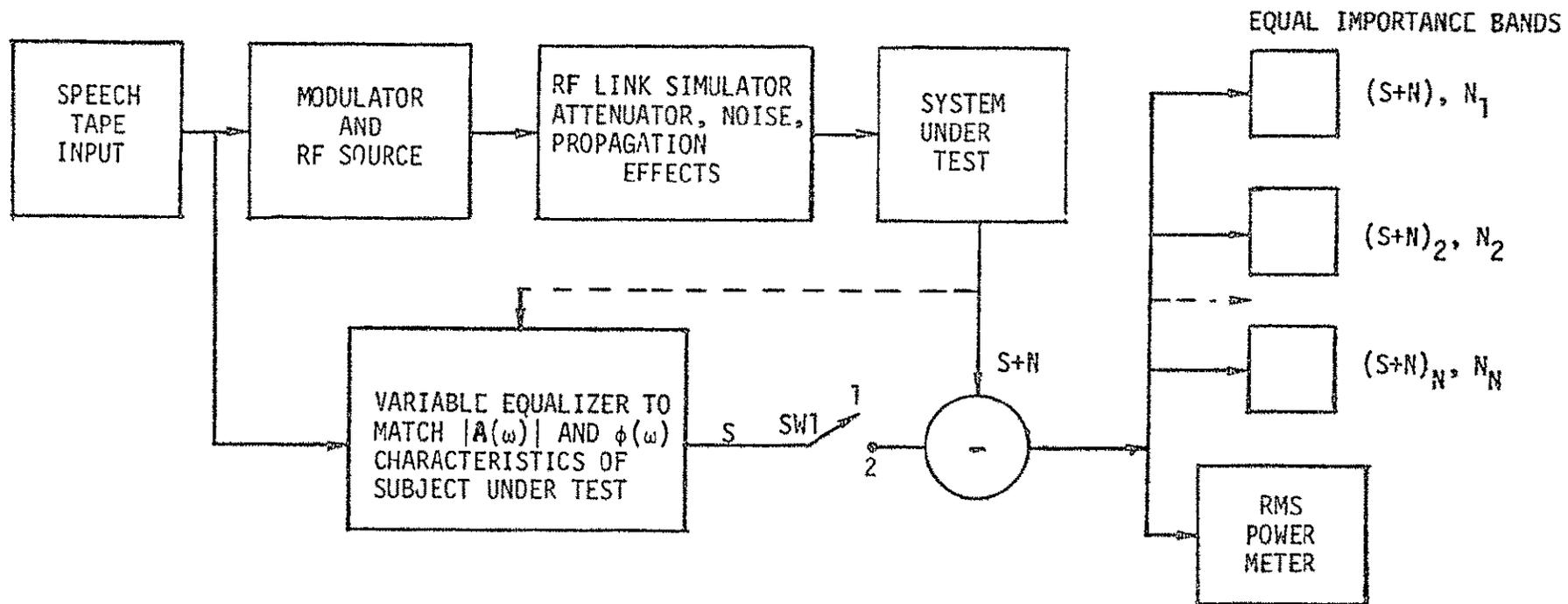
Figure 11. Method A - Determination of
Articulation Index

G   The output of the difference network is applied to an RMS meter
    to read the S+N and N values in order to compute an overall S/N
    ratio

H   The output is also sent in parallel to 20 bandpass filters corre-
    sponding to the equal importance bands of the articulation index
    The S/N ratio of each band is then computed, and the weighting
    functions W_i determined for each band as indicated in Equation (1),
    page 2-5

I   The articulation index (AI) is then computed by Equation (2),
    page 2-5

J   The AI is converted to word intelligibility (WI) by means of an
    empirical relationship determined by subjective testing   Such a
    relationship is controlled by the word list used in preparing
    the input tape   An example of this type of relationship is shown
    in Figure 2

K   The overall S/N ratio computed in Step G could also be converted
    to WI by an appropriately developed empirical relationship as
    shown in Figure 3

3 2 2   Method B - Determination of Articulation Index Using Discrete
        Input Frequencies

Method B is a variation of the method for mechanizing the articulation
index testing in which the input voice signal is replaced by a series of
discrete frequencies centered in each of the "equal importance" frequency
bands   Such a system should be easier to implement since the requirement
for an equalizer to match the input and output would not exist   This method
is suggested by references to "Test Tone SNR" contained in a draft document
of a CCIR study group (Reference 15)   A block diagram is shown in Figure
12

EQUAL IMPORTANCE BANDS

```
┌──────────────┐     ┌──────────────┐     ┌──────────────────┐     ┌──────────┐          ┌────┐
│ SINE WAVE    │     │ MODULATOR    │     │ RF LINK SIMULATOR│     │ SYSTEM   │       ┌─►│    │  (S+N)₁, S₁
│ GENERATORS   │     │ AND          │     │ ATTENUATORS, PRO-│     │ UNDER    │       │  └────┘
│ CENTERED AT  │ ──► │ RF SOURCE    │ ──► │ PAGATION EFFECTS,│ ──► │ TEST     │ ──┬───┤
│ EQUAL IMPOR- │     │              │     │ NOISE            │     │          │    │  ┌────┐
│ TANCE BANDS  │     │              │     │                  │     │          │    ├─►│    │  (S+N)ₙ, Sₙ
└──────────────┘     └──────────────┘     └──────────────────┘     └──────────┘    │  └────┘
                                                                                    │  ┌────┐
                                                                                    │  │RMS │
                                                                                    └─►│POWER│  S+N, S
                                                                                       │METER│
                                                                                       └────┘
```

$(S+N)_1$, S_1

$(S+N)_N$, S_N

S+N, S
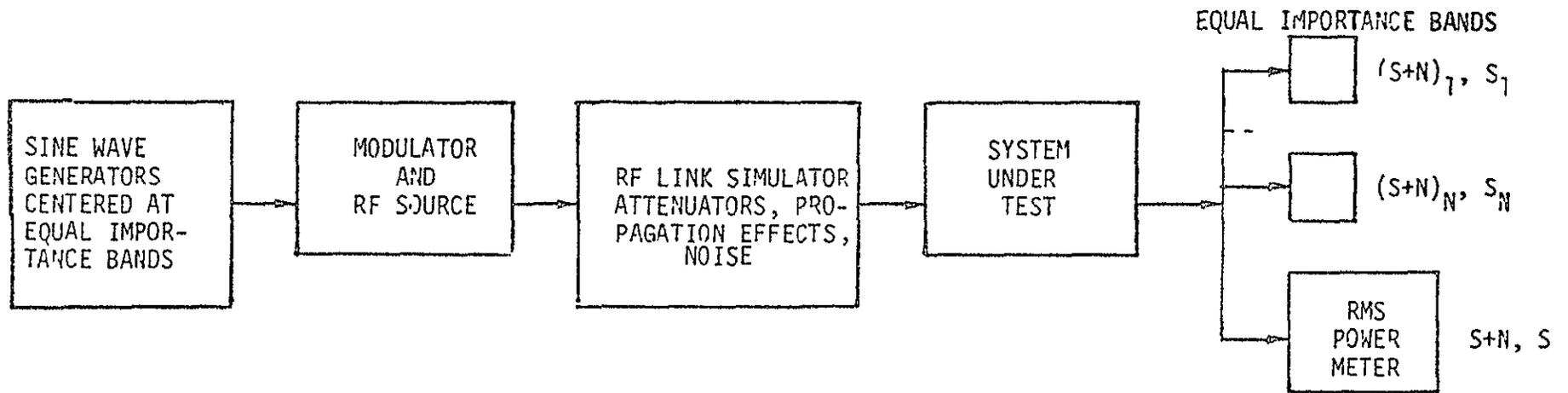
Figure 12. Method B - Determination of Articulation Index
            Using Discrete Input Frequencies

Figure 12

Description of Method

A   The input to the equipment under test consists of a series of single frequencies centered in each of the "equal importance" frequency bands   The amplitudes of the signal frequencies would be adjusted relative to each other in order to conform with a standard speech frequency distribution

B   Modulation, RF source and RF link simulation are determined by the system under test   If the test is an end-to-end test of a complete link, the modulator and RF source would be replaced with components of the system under test

C   A wave analyzer is used to measure the signal power of each of the center frequencies of the equal importance bands   Since the filter characteristics of the analyzer can be very sharp, only a small amount of broad-band system noise will be measured, resulting in measuring only signal power

D   Signal plus noise (S+N) for each of the equal importance bands is measured using filters with the proper response for each band

E.   The S/N ratio of each band is then computed from the signal and signal plus noise measured in steps C and D, and weighting functions derived per Equation (1)

F.   The articulation index (AI) is then computed by Equation (2)

G   The AI can then be converted to word intelligibility (WI) by means of an empirical relationship   Since the WI versus AI characteristic depends upon the length of the word list used, determination of WI would require subjective testing of the system using standard word lists and techniques such as those developed by the U S Standards Association   The WI would be determined using the same average SNR as was used in determining the AI   A relationship between the test signal AI and WI would thus be determined for one value of average SNR

H. Determination of test signal AI for several different average
    SNR's could also be determined, and conversion into WI could be
    accomplished using methods indicated in Step G, resulting in a
    calibration curve.

3.2.3 Method C - Development of Speech-to-Noise Ratio Using
       Analog Measurements

The unique feature of Method C is the method of identifying speech in
the presence of noise, based on the identification of the speech plus noise
and noise only segments of an analog record These analog records are used
to separate and identify the corresponding binary coded decimal (BCD) print-
outs of a digital voltmeter used to integrate the instantaneous squared
waveform of the voice input. The BCD values of speech plus noise and noise
only are then averaged to produce a speech to noise ratio This method was
developed under NASA MSC direction by the Philco Ford Corporation (Reference
13). A block diagram is shown in Figure 13

Description of Method

A This method is based on the ability of recognizing vowel sounds
   which represent approximately 90% of speech power spectrum

B. Standard test tape records scored for word intelligibility will
    be used as input.

C Modulation, RF source and RF link simulation are determined by the
   system under test If the test is an end-to-end test of a complete
   link, the modulator and RF source would be replaced with components
   of the system under test

D The output is converted to an instantaneous squared waveform by
   the true RMS meter, and recorded on a paper recorder.

E Also, a time averaged output over a pre-selected time interval is
   provided by an integrating digital voltmeter

F. The digital voltmeter converts the average power in each sample
    time to BCD and prints out the decimal values. At the same time
    "print commands" from the DVM are recorded on the paper tape along
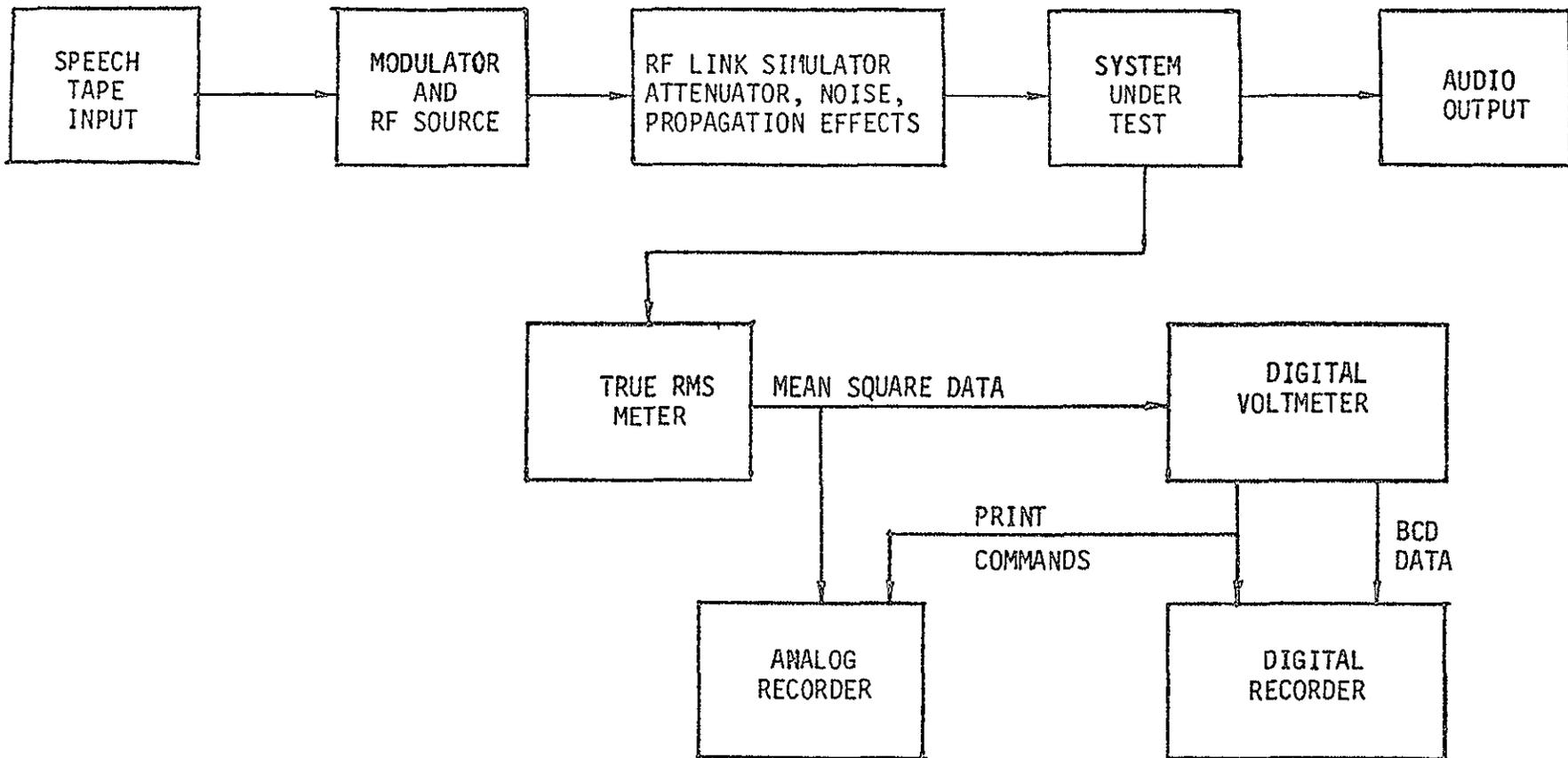
Figure 13. Method C - Development of Speech-to-Noise Ratio
Using Analog Measurements

with the squared analog output. An example of the combined analog
and digital printer output is shown in Figure 14

G. The analog record is used to time correlate the digital printout
of S+N and N. (To distinguish noise samples from speech plus
noise samples.)

H   Speech SNR is then calculated by speech SNR = 10 log $\frac{P1}{P2}$

P1 = average value of speech plus noise

P2 = average value of noise

I. To determine word intelligibility (WI) versus SNR relationship,
several tapes of different quality will be used as input to the
system under test   The output take will be scored by a trained
listening team for WI, and compared to the SNR's.

3 2 4   Method D - Cross Correlation of Input and Output

3 2 4 1   General Discussion

One method of assessing overall system performance is to measure how
well the system input and output functions are correlated (Reference 15)
More specifically, the cross correlation function, $R_{xy}(\tau)$, defined by

$$R_{xy}(\tau) = \int_{o}^{T} X(t + \tau)\, y(t)\, dt \qquad (3)$$

where $x(t)$ is the system input and $y(t)$ is the system output, can be used
as a measure of system performance   The value of $R_{xy}(\tau)$ varies with the
delay, τ, reaching its maximum value when the value of τ approximates the
delay through the system under test. Since the presence of a fixed delay
in a transmission system generally causes no degradation of the transmitted
information, only the maximum value of $R_{xy}(\tau)$ is needed to rate system
performance   If the cross correlation function defined by Equation (3)
is normalized by dividing by the geometric mean of the mean square values
of the input and output signals, the maximum value of the normalized cor-
relation function is confined to the range $-1 \le \rho_{xy}(\tau) \le 1$ The quantity,
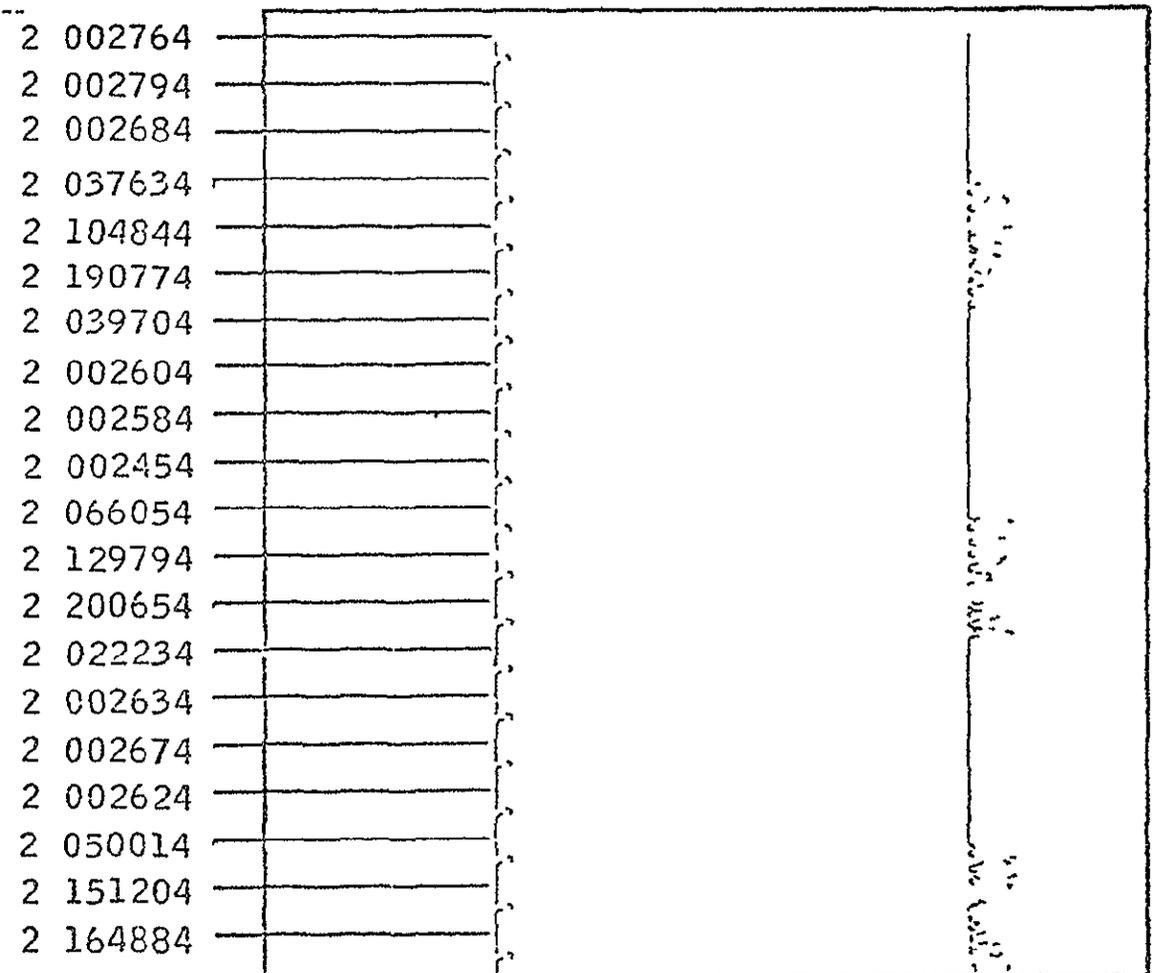
Figure 14 Example of Analog and Digital Printer Output
(Reference 14)

The digital values shown:

2 002764
2 002794
2 002684
2 037634
2 104844
2 190774
2 039704
2 002604
2 002584
2 002454
2 066054
2 129794
2 200654
2 022234
2 002634
2 002674
2 002624
2 050014
2 151204
2 164884

R, needed to specify performance is given by

$$R = |\rho_{xy}(\tau)|_{max} = \frac{|R_{xy}(\tau)|_{max}}{\sqrt{R_{xx}(0)\ R_{yy}(0)}} \tag{4}$$

and is confined to the range $0 \leq R \leq 1$

The quantities $R_{xx}(0)$ and $R_{yy}(0)$ are the autocorrelation functions of the input and output signals evaluated at $\tau = 0$, and are equal to the mean square values of input and output signals respectively If the system under test provides distortionless transmission, the output, $y(t)$, is simple a delayed version of the input, $x(t)$, i.e ,

$$y(t) = Kx(t + \tau)$$

Then from Equation (4)

$$R = \frac{|K \int_0^T x(t+\tau)\ x(t+\tau)\ dt|_{max}}{\sqrt{\left[\int_0^T x(t)\ x(t)\ dt\right]\left[K^2 \int_0^T x(t)\ x(t)\ dt\right]}}$$

$$= \frac{K \int_0^T [x(t + \tau)]^2 dt}{K \int_0^T [x(t)]^2 dt} = 1 \tag{5}$$

If, on the other hand, the system output, $y(t)$, were pure noise when a deterministic $x(t)$ was used as a system input, then

$$y(t) = Kn(t),$$

and

$$R = \frac{K \left| \int_o^T x(t + \tau)\, n(t)\, d\tau \right|_{max}}{K \sqrt{R_{xx}(0)\ R_{nn}(0)}}$$

but

$$\int_o^T x(t + \tau)\, n(t)\, d\tau \rightarrow 0 \quad,\text{ for large } T$$

Consequently, $R \rightarrow 0$.

The block diagram of Figure 15 outlines one method of implementing the normalized correlation coefficient measurement for a receiving system for voice transmission.

3.2 4 2 <u>Description of Method</u>

A   The input tape is composed of phonetically balanced word list scored for percent word intelligibility

B   Modulation, RF source and RF link simulation are determined by the system under test

C   The input tape signal, $x(t)$, and the system output signal, $y(t)$, are fed to the correlation coefficient computer if the test is being made in real time Otherwise, the two signals, $x(t)$ and $y(t)$, are recorded for off-line processing at a later time.

D   The value of the performance parameter, R, is obtained from the computation of the normalized cross correlation function

E   This value of R is used to determine a corresponding value of the Articulation Index, AI, from a calibration curve which has been previously determined by experiment
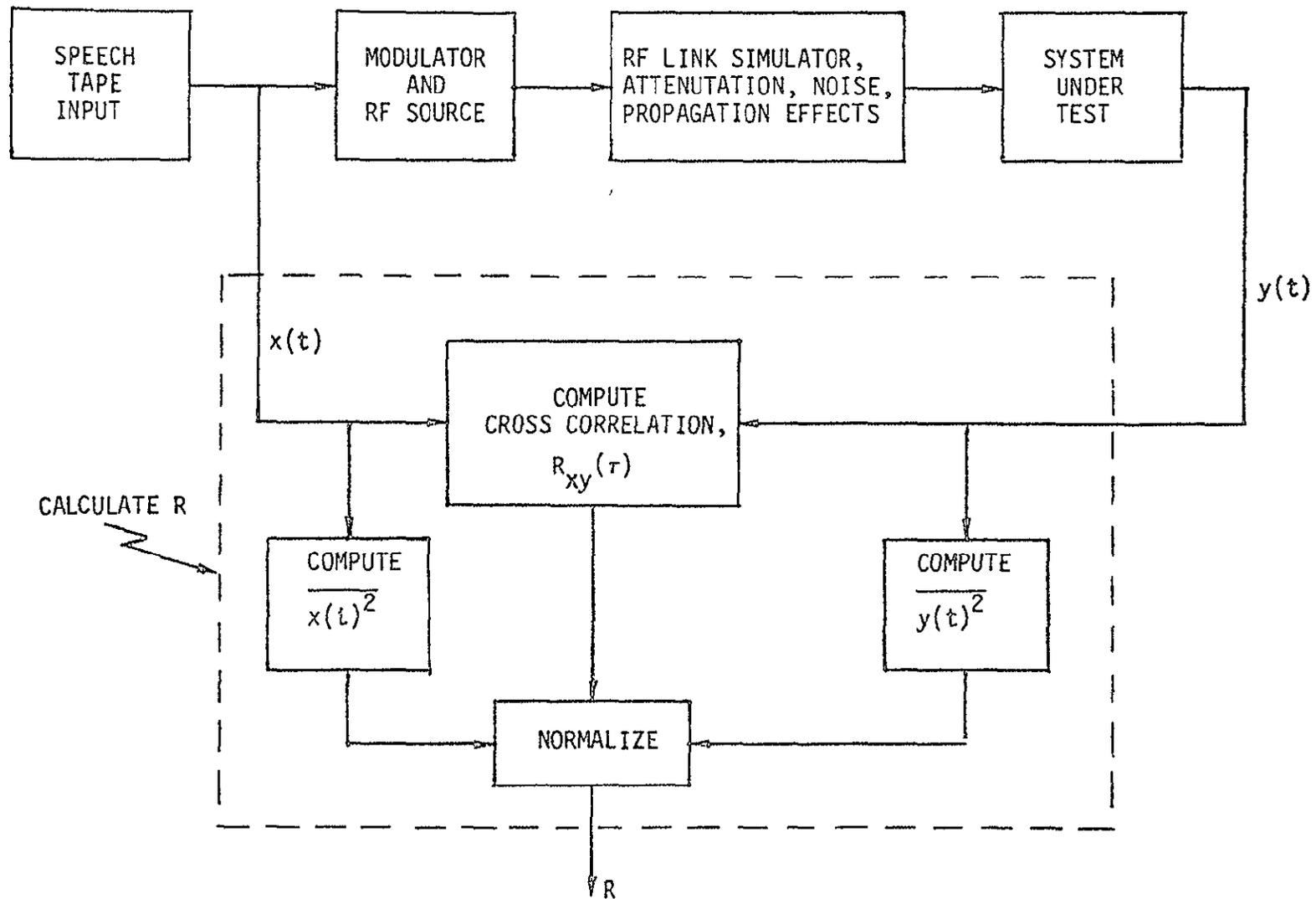
Figure 15. Measurement of the Performance Parameter R

F. The value of AI is converted to percent Word Intelligibility by
   means of the standard calibration curve already determined by
   subjective testing for the particular type and length of word
   list used on the input voice tape.

3 2 5 Method E - Difference of Power Spectra Between Input and Output

This method is based on the fact that the essential intelligence of
speech signals is contained in the short term running power spectrum   The
criterion used therefore is a measure of the mean squared error between
the input and output power spectra of the system under test   The input
and output tapes are converted to digital form, accumulated over a specified
time period, subjected to a Fourier transform routine, and compared in a
difference circuit   The output error is then squared and accumulated at the
end of each test word   The average of the accumulated errors is the evalu-
ation criterion   Figure 16 illustrates a means of mechanizing this method

Description of Method

A. Standard voice test tapes scored for word intelligibility are used
    as the input

B   Modulation, RF source and RF link simulation are determined by the
    system under test   If the test is an end-to-end test of a com-
    plete link, the modulator and RF source would be replaced with
    components of the system under test

C   The tape recording of the output of the system under test is con-
    verted into digital form by an analog-to-digital converter, stored
    in a computer and examined in short time blocks, subjected to a
    Fourier transform routine to compute the power spectra of each
    block, and compared to the power spectra of the input tape pro-
    cessed by a similar routine
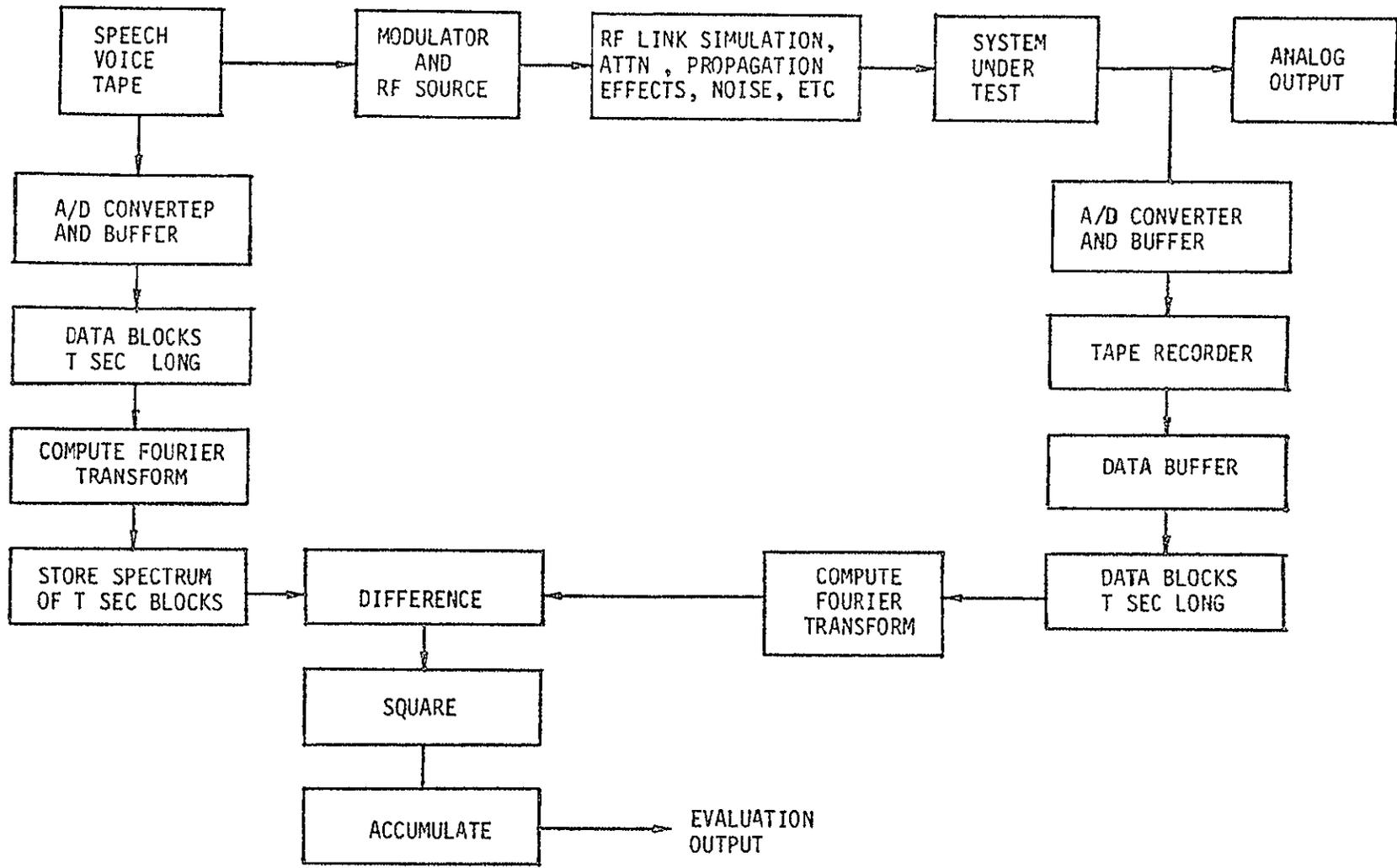
Figure 16   Method E - Mean Squared Error

D   The difference between the input and output spectra is squared
    and accumulated for each test word The average difference is the
    criteria for voice quality

E   To relate to word intelligibility, the output tape would also be
    scored by an experimental test team Several different tapes
    would be used as inputs, and the output word intelligibility for
    each would be plotted as a function of the differences of the
    power spectra, providing a calibration curve

3.2 6   Method F - Digital Method of Determining Speech SNR

Method F is based on the fact that vowel sounds are much longer and
stronger than consonants   It has been estimated that vowels contribute
over 90 percent of the total power spectrum of speech Use can then be
made of this fact to determine speech to noise ratios if monosyllabic test
words (or words spoken so slowly that the syllables can be separated) are
used in preparing input speech tables

This method converts the output of the system under test to digital
form and uses a computer routine to separate speech plus voice from the
noise that occurs between words or syllables To effect this separation
the following assumptions are made

1   The average power of a vowel plus noise waveform will not deviate
    more than 1 dB throughout the duration of the word or syllable
    (100 msec minimum to 200 msec maximum)

2   The average power of the in-between-syllable noise (or in-between-
    word noise) will not deviate more than 1 dB for approximately the
    same time interval as that of a vowel sound

This method was developed under NASA MSC direction by the Philco Ford
Corporation (Reference 14)   A block diagram is shown in Figure 17

Description of Method

A   Standard voice test tapes scored for WI are used as the input

3-16

```
┌──────────┐     ┌──────────┐     ┌──────────────────┐     ┌──────────┐     ┌──────────┐
│ SPEECH   │     │ MODULATOR│     │ RF LINK SIMULATOR,│     │ SYSTEM   │     │ ANALOG   │
│ TAPE     │ ──► │ AND      │ ──► │ ATTN, PROPAGATION │ ──► │ UNDER    │ ──► │ OUTPUT   │
│ INPUT    │     │ RF SOURCE│     │ EFFECTS, NOISE,   │     │ TEST     │     │          │
│          │     │          │     │       ETC         │     │          │     │          │
└──────────┘     └──────────┘     └──────────────────┘     └──────────┘     └──────────┘
                                                                 │
                                                                 ▼
                                                           ┌──────────────┐
                                                           │ A/D CONVEPTER│
                                                           └──────────────┘
                                                                 │
                                                                 ▼
                                                           ┌──────────────┐
                                                           │ COMPUTER     │
                                                           └──────────────┘
                                                                 │
                                                                 ▼
                                                           ┌──────────────┐
                                                           │ LINE PRINTER │
                                                           └──────────────┘
```
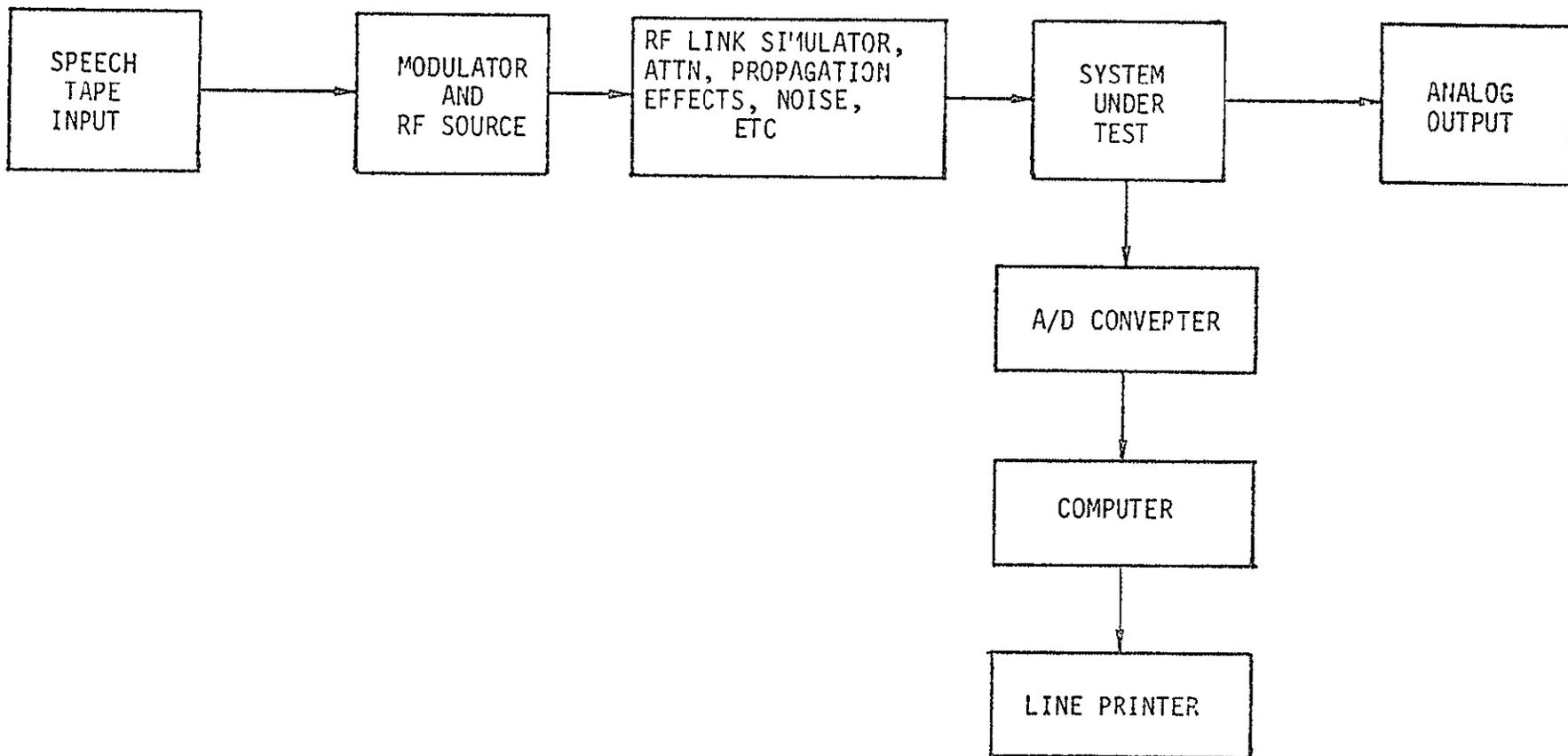
Figure 17.   Method F - Digital Method of
             Determining Speech SNR

B Modulation, RF source and RF link simulation are determined by the system under test If the test is an end-to-end test of a complete link, the modulator and RF source would be replaced with components of the system under test

C After analog to digital conversion, the output is applied to a computer with a program designed to mechanize solution of the problem and provide the output on a line printer

D The computer program is based on

1 Use of 20 msec as the measurement interval, with the average power in at least three consecutive intervals being compared to be within 1 dB, (each 20 msec interval contains 400 samples at sample rate of 20K)

2 Three or more consecutive intervals are averaged and placed in storage until 50 consecutive intervals have been accumulated

3 The logic assumes that the smallest value of average power for three or more consecutive intervals (which agree within 1 dB) represents noise and this value is taken as a reference level

4 All other values of (three or more consecutive intervals) are compared to this reference level

5 If a value compares with 1 dB of the reference value, it is considered noise for the purpose of computation

6 If a value is 3 dB or greater than the reference value, it is considered S + N for the purpose of computation

7 Values between 1 dB and 3 dB or the reference value are ignored

8 The mean is calculated for all values of N and for all values of S + N, thus providing the basis for speech signal-to-noise

ratio (SPNR) computation for each are second interval·

$$SPNR = \frac{(S+N) - N}{N}$$

Thus, the minimum SPNR is 0 dB and occurs when S + N = 2N

E   The value of SPNR calculated above are converted to word intelli-
    gibility by producing a system output analog tape for scoring by
    trained observers   Several input tapes of different quality would
    be used to produce a SPNR versus WI calibration

3 2.7   Method G - Bit-by-Bit Comparison of Input and Output Tapes

This method is simple in concept in that it merely takes the input
and output of the system under test, converts them into digital form and
compares them bit-by-bit to determine the bit-error-rate   BER versus word
intelligibility calibration would have to be made by subjective testing of
the analog output of the system under test.   Figure 18 is a block diagram
illustrating this method

Description of Method

A   As a system input, standard voice test tapes of scored for word
    intelligibility (WI) will be used

B   Modulation, RF source and RF link simulation are determined by
    the system under test   If test is end-to-end, the modulator and
    RF source would be replaced with components of the system under
    test

C. The output of system under test is converted into digital form by
    an A/D converter, in addition, an analog tape output is provided
    for comparison

D   The input is also fed through a delay calibrated to match the system
    delay to an identical A/D converter.

E   The output of the two A/D converters are compared on a bit-by-bit
    basis and a bit error rate (BER) is calculated
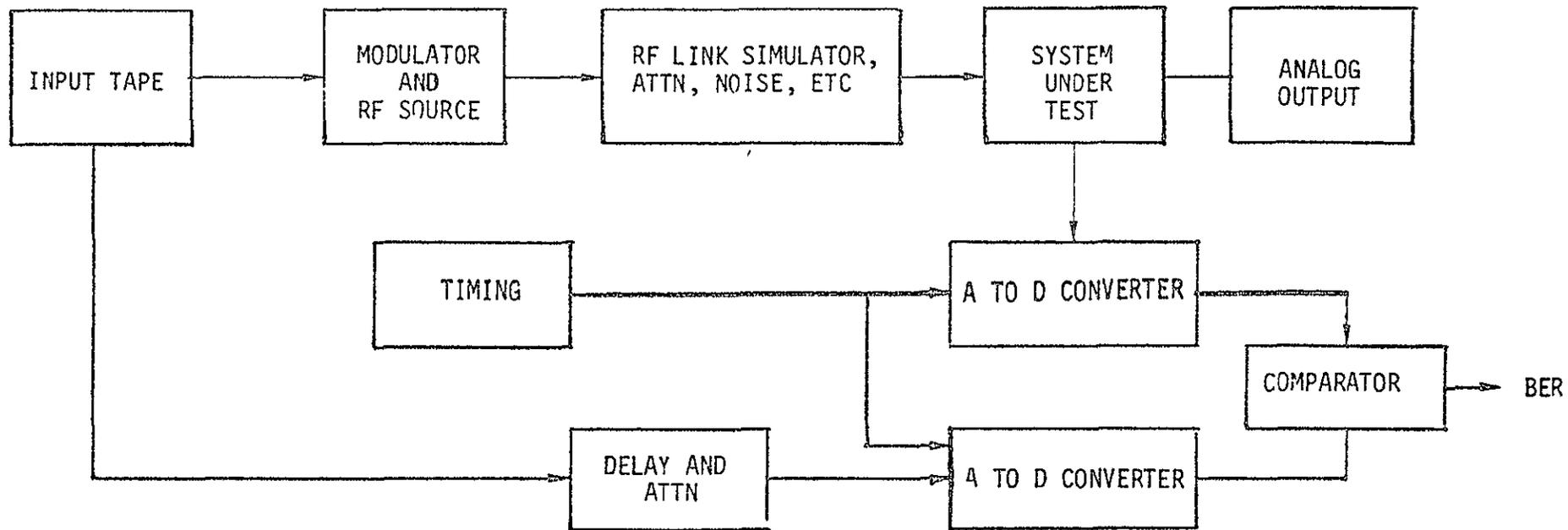
VOICE CRITERIA - METHOD G

BIT-BY-BIT COMPARISON

Figure 18

, F. To establish a calibration of BER versus WI, the output analog
tape would be scored by a trained observer team Input tapes of
different qualities would be used to produce a BER-WI calibration
curve.

3 3 VIDEO SYSTEMS CRITERIA

3 3 1 Method A - Video Signal to Noise Ratio Measurement Using a Weighted Noise Concept

Use of a noise weighting scheme in determining picture signal-to-noise
ratios (SNR) is based on the fact that noise in the lower end of the video
spectrum has a greater effect on picture quality than noise at the upper
end of the spectrum Methods using noise weighting have been investigated
by several groups of researchers such as the International Radio Consulta-
tive Committee (CCIR), the Electronic Industries Association (EIA), Bell
Telephone Allocation Study Organization (TASO), and the United States
Standards Association This method is described in more detail in Section
2 3 Since the noise weighting curve is an experimentally determined
relationship describing relative video picture degradation as a function
of noise frequency, use of the noise weighting curve should be applicable
to systems which have non-flat video noise spectrums as well as to those
which have white noise For example, the noise weighting function should
be applicable to frequency modulated television systems which result in
parabolic noise Figure 19 is block diagram of Figure 19 is a block diagram
of a system to measure weighted picture SNR

Description of Method

A A video tape or a color slide scanner and selected color slides
will be used to provide video input to the system under test A
noise generator capable of providing a flat noise spectrum over
the video bandwidth may be used to inject noise into the system
either at the input or as part of the link simulator

B Modulation, RF source and RF link simulation are determined by
the system under test If the test is an end-to-end test of a
complete link, the modulator and RF source would be replaced with
components of the system under test
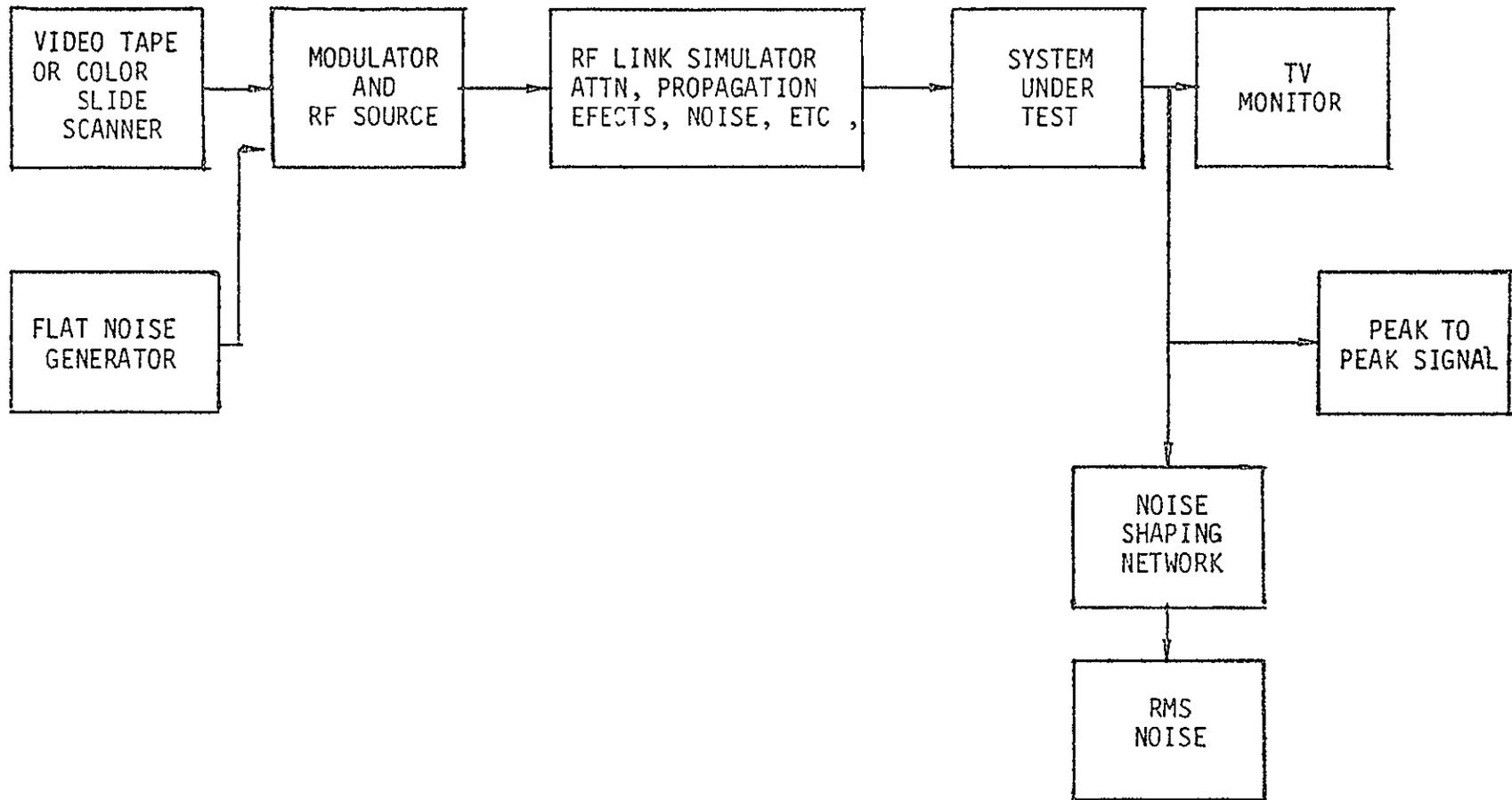
Figure 19. Method A - Video Signal to Noise Ratio Measurement
Using a Weighted Noise Concept

C   A noise shaping network at the output will provide noise weighting
    in accordance with the noise weighting curve provided by the United
    States Standards Association   This curve is shown in Figure 8
    The weighted noise will be measured by a true RMS voltmeter when
    no video signal is provided to the input of the system under test

D   With the video signal applied, the white to blank video signal
    will be measured

E. The weighted picture SNR will be calculated·

$$S/N = \left(\frac{\text{blank to white video voltage}}{\text{weighted RMS voltage of video noise}} \right)^2$$

F   At the same time, an observer team will make an assessment of the
    quality of the output picture on the TV monitor, using a standard
    5 or 6 point Bell Laboratories, CCIR or TASO scale. Use could be
    made of the curves shown in Figure 9, but improvements made in TV
    equipment since the time the curves were taken could make this
    curve obsolete

G   Several combinations of video signal level input and noise input
    (either at the input or as part of the link simulation) will be
    measured and scored by the observer team, resulting in a calibra-
    tion curve of picture SNR versus picture quality.

3 3 2 Method B - Cross Correlation of Input and Output

Description of Method

      This method is essentially the same as that described for Method
D, Voice Criteria, with the substitution of a standard video tape or
the output of a slide scanner used in place of the voice tape as input
to the system under test   The analog-to-digital conversion of the
input x(t) and output y(t) to the correlation computer would operate
at a higher data rate because of the video bandwidth, but the princi-
ples and method would be the same as those described for the voide
system

3 3.3 Method C - Equal Importance Frequency Bands

Description of Method

This method is hypothetical and is based on the fact that in the weighted noise concept of measuring video signal-to-noise ratios, bands of noise centered at different frequencies in the video band cause equal subjective interference effects when the applied through a noise weighting network (Reference 7) The frequency composition of a typical series of noise bands is shown in Figure 20 When the noise amplitudes in these bands were weighted in accordance with the standard of the U. S Standards Association, equal SNR's caused equal subjective effects as judged by a panel of observers Additionally, a more or less linear relationship of a picture quality rating scale and the weighted noise level in dB appears to exist Thus it is postulated that it might be possible to divide the video frequency bands into a number of "equal importance" frequency bands which could be used as a basis for establishing video criteria in a manner similar to that used in calculating the articulation index of voice systems

The method used in such a system would be similar to that described in Method A, Voice Criteria

3.3 4 Method D - Mean Squared Error of Input and Output Spectra

Description of Method

This method would be essentially the same as that described for Method E - Digital Voice Criteria, with the substitution of a standard video tape or the output of a slide scanner in place of the voice tape as input to the system under test The A to D converters would require a larger number of quantizing levels (128 levels, or 7 bits has been suggested as adequate for picture information encoding, Reference 16) and higher data rates, but the technique would be the same The analog video output would be scored for quality by an observer team to provide a calibration relationship to the difference of the mean squared errors
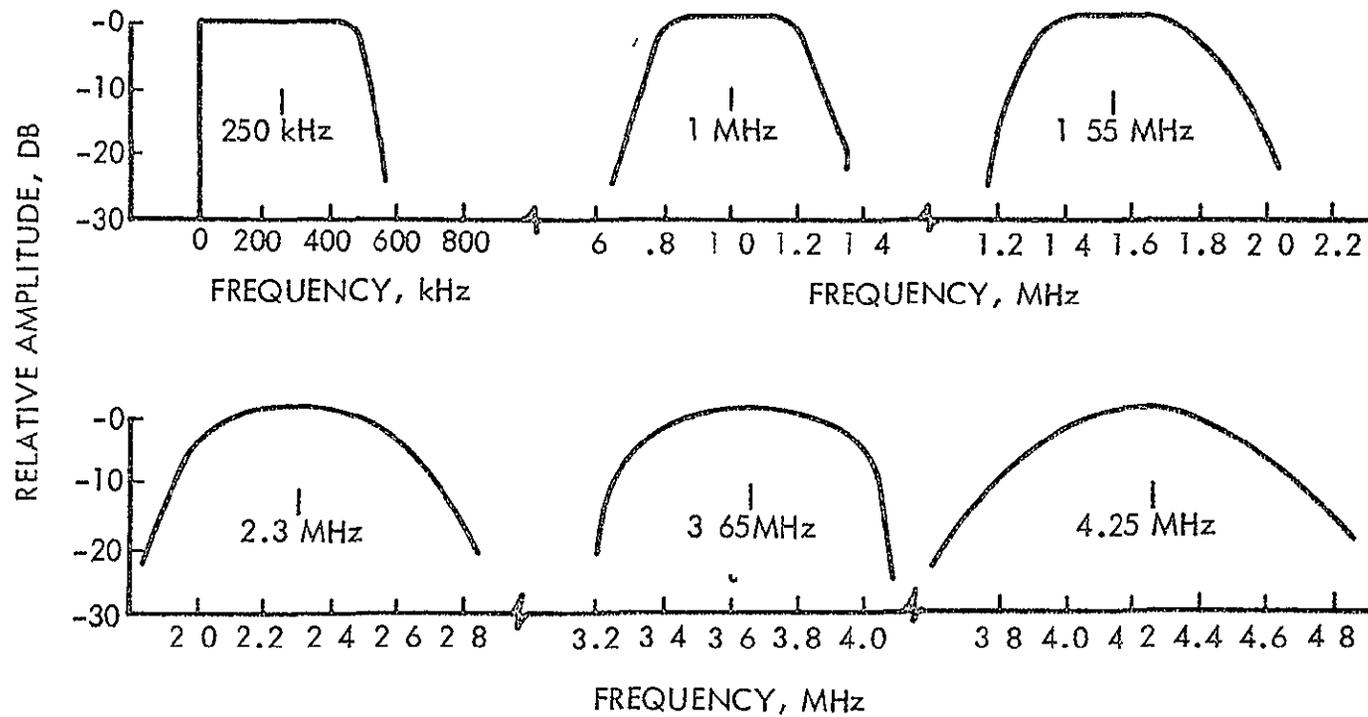
Figure 20   Frequency Composition of 500-kHz Bands of Random Interference
Centered at Various Video Frequencies

3 3.5 Method E - Bit-by-Bit Comparison of Input and Output

Description of Method

This method would be similar to Voice Criteria - Method G   A standard video tape or the output of a slide scanner would provide the input   The A to D converters and digital comparators would be required to operate at a much higher data rate   The analog video output would be scored for quality by an observer team in order to provide a calibration relationship of picture quality versus BER

3 4 DIGITAL DATA SYSTEM CRITERIA

3 4.1 Method A - Bit-by-Bit Comparison

Bit-by-bit comparison of digital tapes of the input and output of a system under test is conceptually one of the most simple methods of determining bit error rate. Timing and synchronization of the input and output can be a problem   For very low error rates, the counting time may be appreciable   A block diagram is shown in Figure 21

Description of Method

A   The input to the system under test is digital test tape of known message content.

B   Modulation, RF source and RF link simulation are determined by the system under test   If the test is an end-to-end test of a complete link, the modulator and RF source would be replaced with components of the system under test.

C. The output of the system under test is fed to a comparator where it is compared to the input bit stream, after the input bit stream has been corrected for system delay

D   The comparator produces a BER directly by counting the errors in a specified length of time   A problem with this method is that if the BER is very low, an unacceptably long time may be required to count enough errors to give a reliable estimate of the actual error rate
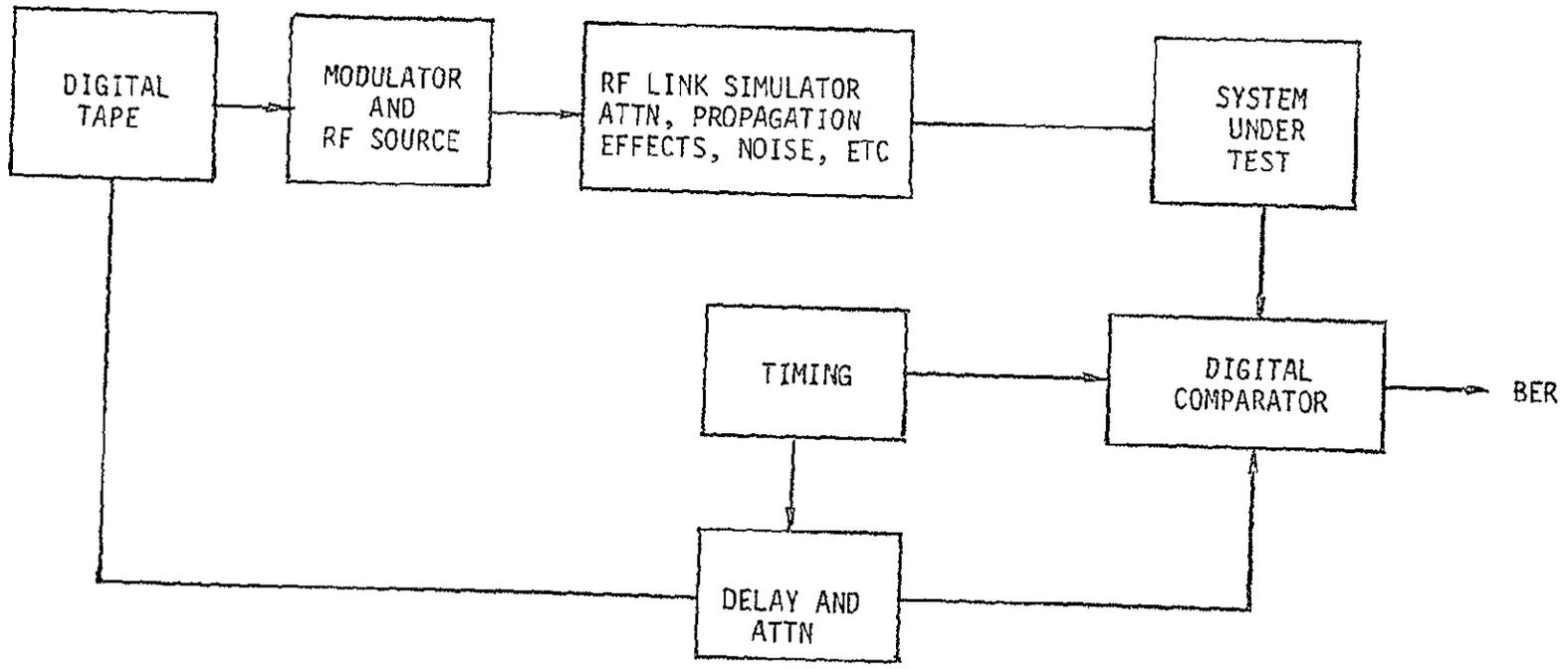
Figure 21   Method A - BER by Bit-by-Bit Comparison

3 4 2 Method B - Pseudo-Error Extrapolation

3.4 2.1 General Discussion

In an effort to overcome the problem of long counting time which may occur in conventional bit-by-bit comparison of input and output data streams when the error rate is very low, the technique of computing pseudo error rates which are much larger than the actual error rate, has been developed (Reference 11) This method, illustrated in Figure 22, creates large pseudo error rates by biasing modified one-zero decision circuits in favor of the incorrect decision, and is characterized by the following features

1) The pseudo error rates are generated by use of modified decision thresholds in the "one" and "zero" channels

2) A method of estimating the pseudo error rates corresponding to two or more modified decision thresholds.

3) Two or more estimated pseudo error rates based on different decision thresholds are used to generate a function of pseudo error rates versus a parameter representing the modified decision thresholds

4) This function is extrapolated to a point where the decision threshold parameter corresponds to that of the actual decision threshold Thus, at this point the estimated pseudo error rate equals the estimated actual error rate

The principle of operation of a device designed to produce a pseudo-error rate (P_p) is based on the following observation "for a given type of modulation and given form of probability distribution of the noise and fading processes, it is possible to define a threshold parameter K such that the logarithm of the pseudo-error rate P_p is a linear function of K for a wide range of values of P_p The linear portion of this curve, when extended to K=0, coincides with the logarithm of the actual error rate Thus, by measuring P_p for two values of the parameter K and linearly extrapolating through these two points to the value K=0, one obtains an estimate of the logarithm of the actual error rate" (Reference 11) This estimate
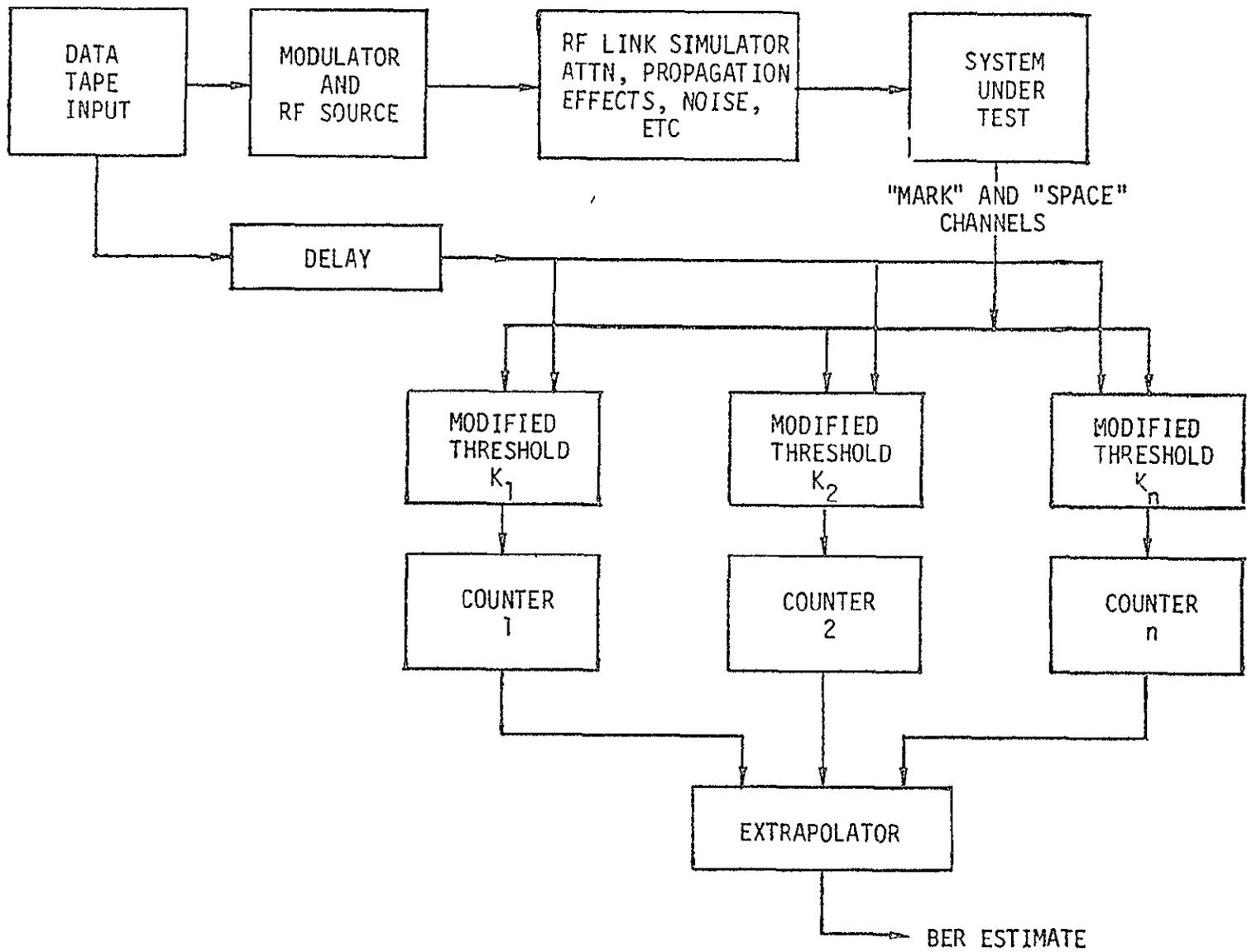
Figure 22   Method B - Pseudo-Error Rate Extrapolation

of actual error rate is valid for the particular SNR used in calculating the P_p versus K curve Thus, if it is desired to produce an estimated BER versus SNR curve the procedure would have to be repeated to determine the estimated actual BER for the number of SNR values desired

Figure 23 shows the graph (A) of the logarithm of the actual receiver error rate P_e as function of signal-to-noise ratio (R) for some propagation criteria, plotted with curve (B) of the logarithm of the pseudo error rate P_p versus the parameter K for a particular value of received signal-to-noise ratio, R_0 The point where the linear extrapolation of curve (B) intersects the R_0 ordinate is also the point when the curve (A) intersects the R_0 ordinate Curve (A) is included to illustrate the principle involved In practice, of course, curve (A) is the desired result The log P_p curve (B) would be extrapolated from the points (x) and (y) representing the log P_p values for threshold parameter values (K_1) and (K_2) respectively, which would result from the condition where the value of SNR is (R_0) Where the linear extrapolation of curve (B) intersects the value of K=0 represents the value of the estimated BER for the value of SNR (R_0) for which the points (x) and (y) were calculated This process would have to be repeated for each SNR value for which it is desired to estimate the actual value of BER

Since this method of computing pseudo-errors depends on modified thresholds in the "one" and "zero" channels, access to these points prior to the threshold detector must be made available in the system under test

3 4 2 2 Description of Method

A The input to the system under test is a digital test tape of known message content

B Modulation, RF source and RF link simulation are determined by the system under test If the test is an end-to-end test of a complete link, the modulator and RF source would be replaced with components of the system under test.
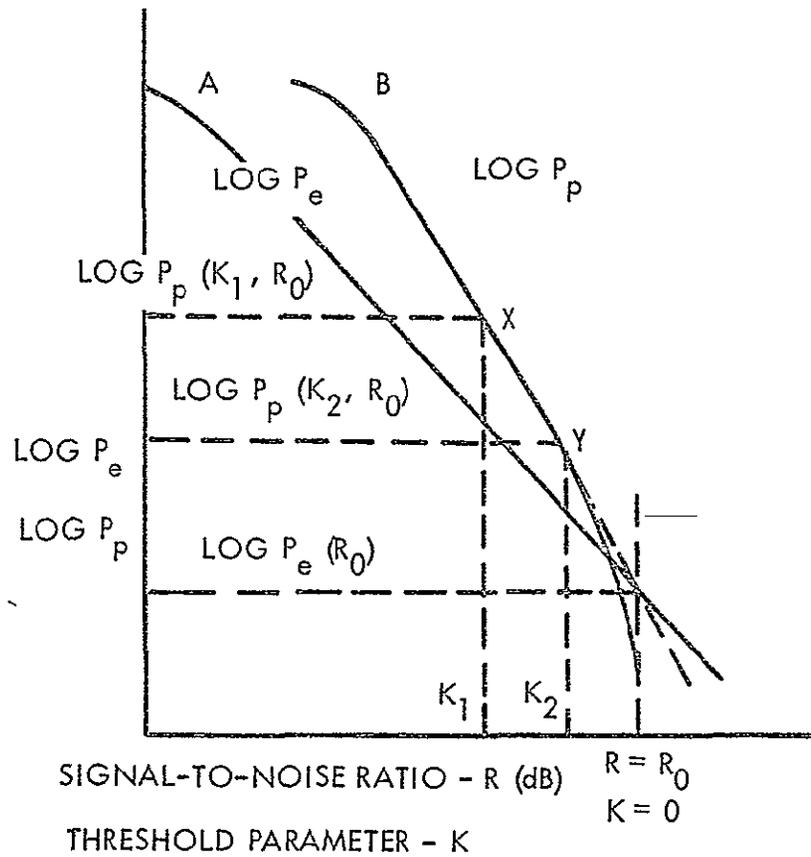
Figure 23   Plot of Pseudo Error Rate (P_p) Versus Threshold
Parameter K Compared to Plot of Actual Error
Rate (P_e) Versus SNR

C. The "one" and "zero" outputs (prior to the threshold detection) of the system under test are fed to a series of modified decision circuits where they are compared with the input data stream in such a manner that the pseudo error rates generated are larger than the actual error rate of the output

D. The pseudo error rates are counted and fed to an extrapolator where a linear extrapolation of the pseudo errors versus their respective threshold parameters K, K_2, K_n is made to extend to the point $K=0$ (the point at which the modified threshold is equal to the actual threshold in the system under test) At this point, the estimated pseudo error rate is equal to the actual estimated error rate

E. Since the pseudo error rates are larger than the actual error rates, the time to count the estimated error is much shorter than that required by a bit-by-bit counting process

3 5 COMPARISON OF CANDIDATE SYSTEMS

3 5 1 Rating of Comparison Parameters

In order to determine the relative value of each of the candidate criteria systems, some scheme of numerical rating must be used The method used here is to assign a rating number for each comparison parameter for each of the candidate systems This number ranges from 1 to 5, with number 1 representing the best However, the numbers are not exclusive That is, if for any particular parameter, such as Precision of Test Data, it is felt that two of the candidate systems result in about the same precision of data, each is assigned the same numerical rating These rating numbers, when multiplied by the weighting values discussed in Section 3 5 2 yield a value for each parameter for each of the candidate systems These parameter values, when added for each candidate system, give an indication of the overall rating for each system, with the system resulting in the lowest total value being considered the best

Each parameter rating was assigned after consideration of the pertinent factors for each parameter for each of the candidate systems One of the

basic difficulties in the rating scheme is the definition of the comparison parameters. For the purpose of this report the parameter "Precision of Test Data" is used more or less synonymously with accuracy. In assigning values to Parameter One, Precision of Test Data, each factor listed was considered for accuracy, and an "average" accuracy of each system was determined. Those systems with the best "average" accuracy were assigned a rating of 1, etc

The parameter "Conciseness of Results" presents something of a problem of interpretation. Since all of the candidate systems provide a number value output for a given input, it could be said that they were equally concise. The values arrived at in this report, however, are based on the relative amount and complexity of the computations required to achieve the final value for each of the candidate systems

The parameters "Data Reduction Required" and "Ease of Simulation" are interdependent and are nearly redundant in concept. As used herein, "Data Reduction" is taken to indicate not only the total amount of data computation involved in a given system, but also the amount of computation required by the operator after the process has been completed. It can be seen, then, that "Data Reduction Required" and "Ease of Simulation" are somewhat reciprocal - a system with a low rating score for required data reduction could be expected to have a relatively high rating score for ease of simulation.

"Relation to Actual System Performance" is the most difficult of comparison parameters to rate for some of the candidate systems. Most of the systems require calibration by subjective testing to establish a relationship between the quantity derived as a result of the test and the desired result - percent word intelligibility or picture quality. Since some of the candidate systems have not been mechanized, the question of whether there is a monotonic relationship between the quantity derived from the test and a subjective evaluation can only be determined by test

3 5.2 Weighting of Comparison Parameters

After first determining the parameters to be used in evaluating performance criteria, a second question to be considered is that of the relative weighting of each of the parameters Do all of them affect the overall value of performance equally, or are some of them more important than others? The answer to this question depends upon those who are using the criteria to evaluate the performance of a system

In determining a weighting system, the precision of the test data appears an obvious choice as the most, or one of the most, important parameters and was thus assigned a value of 1 Relationship to actual system performance seems almost as important as precision of data and was also assigned a value of 1 Ease of simulation and the amount of data reduction required were determined to be of about equal importance, but were judged to be of less critical nature, and were assigned values of 2 The last parameter, "Conciseness of Results", was assigned a value of 3, not so much because it was felt to be of less importance than the others, but because the very nature of the candidate criteria systems is such that the results tend to represent averages of measured and computed values In addition, the results of criteria systems for voice and video quality must be related to subjective evaluation

This arbitrary numerical weighting, which is listed below, has been used in Table 3, which lists numerical values of the parameters for each of the candidate systems, as well as the total rating for each system

Parameter	Weighting
Precision of Data	1
Ease of Simulation	2
Data Reduction Required	2
Relation to Actual System Performance	1
Conciseness of Results	3

Criteria Evaluation Weighted Parameters

Voice Systems

		Precision of Data	Conciseness of Results	Data Reduction Required	Relation to Actual System	Ease of Simulation	Total Values	Overall Ranking
A	AI-Equal Importance Bands	1	3	4	1	6	15	3
B	AI-Discrete Frequency	1	3	4	2	2	12	1
C	Speech SNR-Analog	2	6	6	3	6	23	5
D	Cross Correlation	2	6	6	5	8	27	7
E	Mean Squared Error	2	9	2	3	8	24	6
F	Speech SNR-Digital	2	6	2	3	8	21	4
G	Bit-by-Bit Comparison	1	3	2	4	4	14	2

Video Systems

		Precision of Data	Conciseness of Results	Data Reduction Required	Relation to Actual System	Ease of Simulation	Total Values	Overall Ranking
A	Picture SNR	1	3	2	1	2	9	1
B	Equal Importance Bands	3	6	6	5	6	26	5
C	Cross Correlation	2	6	6	5	6	25	4
D	Mean Squared Error	3	9	6	3	4	25	3
E	Bit-by-Bit Comparison	2	6	6	3	6	23	2

Digital Data Systems

		Precision of Data	Conciseness of Results	Data Reduction Required	Relation to Actual System	Ease of Simulation	Total Values	Overall Ranking
A	Bit-by-Bit Comparison	1	3	2	1	2	9	1
B	Pseudo Error Extrapolation	2	3	2	3	4	14	2

Table 3

Another weighting system, or none at all (assuming all the parameters to have equal importance) could be used, depending upon the needs of the users of the performance criteria The intention here is to indicate how such a weighting scheme, when combined with a rating for each parameter assigned to each of the candidate systems can assist in determining the overall ranking of a particular candidate system Table 4 lists the unweighted values, assuming that each of the parameters has equal weight

3 6 NEW CRITERIA SYSTEMS SUMMARY AND CONCLUSIONS

3 6.1 General Considerations

Some of the difficulties pertaining to the assignment of numerical ranking of the different criteria systems which have been pointed out in this section are

1. The difficulty associated with defining the parameters

2 Use of a weighting system for the comparison parameters, and the weighting values assigned to each parameter if used

3 Rating each system for each parameter

4 The uncertainties associated with untried methods

One problem not treated in the section is that of differentiating between criteria and methods used to test a system to meet that criteria The approach taken in this report is to consider criteria systems or methods This leads to some duplication as regards criteria - for instance, the two methods described to achieve voice articulation index Since the two methods result in different ratings, it is felt that this approach is of value

An additional factor not considered in the final selection of a per-formance criteria is the ease with which the criteria can be used by the systems or equipment designer in developing the design of the system For example, the performance of a portion of a system might be specified in terms of a criterion which is accurately related to system performance,

Criteria Evaluation Unweighted Parameters

Voice Systems

		Precision of Data	Conciseness of Results	Data Reduction Required	Relation to Actual System	Ease of Simulation	Total Values	Overall Ranking
A	AI-Equal Importance Bands	1	1	2	1	3	8	2
B	AI-Discrete Frequency	1	1	2	2	1	17	1
C	Speech SNR-Analog	2	2	3	3	3	13	4
D	Cross Correlation	2	2	3	5	4	16	6
E	Mean Squared Error	2	3	1	3	4	13	5
F	Speech SNR-Digital	2	2	1	3	4	12	4
G	Bit-by-Bit Comparison	1	1	1	4	2	9	3

Video Systems

		Precision of Data	Conciseness of Results	Data Reduction Required	Relation to Actual System	Ease of Simulation	Total Values	Overall Ranking
A	Picture SNR	1	1	1	1	1	5	1
B	Equal Importance Bands	3	2	3	5	3	16	5
C	Cross Correlation	2	2	3	5	3	15	4
D	Mean Squared Error	3	3	3	3	2	14	3
E	Bit-by-Bit Comparison	2	2	3	3	3	13	2

Digital Data Systems

		Precision of Data	Conciseness of Results	Data Reduction Required	Relation to Actual System	Ease of Simulation	Total Values	Overall Ranking
A	Bit-by-Bit Comparison	1	1	1	1	1	5	1
B	Pseudo Error Extrapolation	2	1	1	3	2	9	2

Table 4

but which is very difficult for the designer to compute. Although this computational difficulty is not an overriding consideration in criteria selection, it could be considered in assigning a weight to the trade-off parameter "Ease of Simulation "

3 6 2 Voice Systems

Based on both the unweighted and weighted parameters, calculation of articulation index using discrete frequencies centered in the equal importance bands appears to be the best method for specifying voice performance This ranking is different from that published in TRW report No 17618-H174-RO-OO, "New Criteria Development", in that the original ratings showed Method A using voice signals to be superior   Further consideration of the relative difficulty of mechanization and relation to actual system performance resulted in changes in these parameters and in the total scores.

References to Tables 3 and 4 reveals the composition of the ratings of the other systems   It is of interest to note that the weighted and unweighted overall ranks are quite similar   The important feature of the tables is that it assists users with different requirements to determine which of the systems would be more suited to his needs   For instance, if the amount of data reduction required were not of prime importance to a particular user, he could downgrade or ignore this particular parameter and re-compute the total for each system, thus arriving at a rating suited to his requirements

3 6 3 Video Systems

The picture SNR method using a standard noise weighting is the clear choice based on both the weighted and unweighted parameter systems   The vact that this scheme has had considerable proven experience undoubtedly affected the results

3 6 4 Data Systems

There is really only one choice for the digital systems criteria - bit error rate based on comparison of input and output   Method B is actually a sub-method, and under the circumstances of very low BER conditions could be the number one choice

4. CRITERIA TEST MECHANIZATION

4.1 GENERAL DISCUSSION

Section 4 presents several methods of mechanizing tests designed to determine compliance with performance criteria discussed in previous sections The criteria systems discussed in this section are determination of voice articulation index (AI) using discrete frequency inputs, video picture signal-to-noise measurements using weighted noise, and use of pseudo bit error rate measurements for digital data

4.2 DETERMINATION OF ARTICULATION INDEX USING DISCRETE INPUT FREQUENCIES

4 2.1 Input Signals

As discussed in Section 1, the input to the system under test is a magnetic tape recording consisting of calibration signals to check out the standardized test equipment and test signals to be applied to the system under test

The test signal inputs will consist of a set of signal frequencies centered in each of the equal importance bands listed in Table 1. These "band center" frequencies are tabulated in Table 5. It will be noted that the frequencies chosen for Bands 2, 8, 10, 12, and 20 are not exactly midband These frequencies have been slightly offset to minimize the effects of possible harmonics of the test frequencies resulting from non-linear system operation. This precaution is taken because of the relatively low level of the high test frequencies. The frequencies chosen insure that possible harmonics up to and including the seventh, are separated from any test frequency and any other harmonics by at least 10 Hz. The power contained in each frequency is based on the normal voice spectrum illustrated in Figure 24 For each band, the power was determined by first calculating the average power level in each band by determining the average of the band edge power levels. This average power level was multiplied by the bandwidth in Hz. Finally, the power in each band was normalized by dividing the product by the sum of all the products in the 20 bands, as shown in Equation 6. The resulting relative power levels are listed in Table 5.

TABLE 5
Discrete Frequencies for Equal Importance Bands

| Articulation Band Number	Band Center Frequency in Hz	Normalized Power Level
1	265	0.2065
2	375	0.1758
3	495	0 1960
4	630	0.1440
5	770	0 0925
6	920	0 0628
7	1075	0.0346
8	1235	0 0235
9	1395	0 0161
10	1560	0.0114
11	1745	0.00743
12	1915	0 00611
13	2130	0 00567
14	2370	0.00528
15	2660	0 00505
16	3010	0.00439
17	3425	0 00329
18	3950	0.001945
19	4650	0 000763
20	5560	0 000475
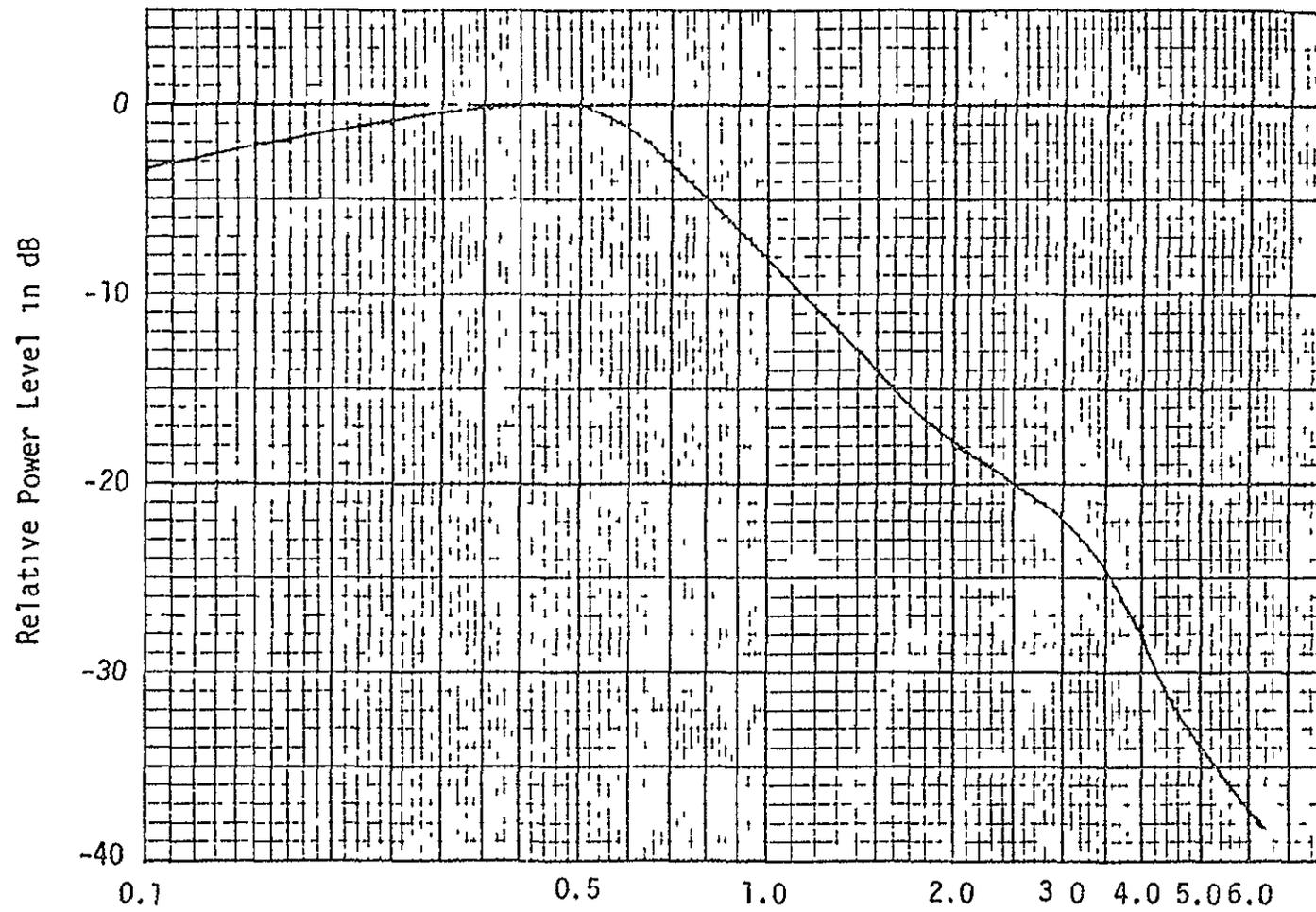
Figure 24   Frequency in kHz Normal Voice Spectrum

$$P_i = \frac{\frac{P_2 + P_1}{2} \times BW_i}{\sum_{n=1}^{20} P_i} \qquad (6)$$

where P_1 and P_2 are the power levels at the band edge frequencies, and BW_i is the bandwidth in Hz of band i.

As shown in Figure 25, the input test frequencies could be recorded simultaneously by using separate oscillators set at the band center frequencies with the amplitudes adjusted in accordance with Table 5

4.2.2 Output Signal Mechanization

There are a number of methods of mechanizing the calculation of the AI from the output of the system under test, varying from the simple manual type operation shown in Figure 26 to more automated mechanization

In Figure 26 the signal and noise power is measured at the output of each band filter, and a wave analyzer is used to measure the signal power of each center band frequency   The output of the band filters is signal plus noise, but since the filter characteristics of the wave analyzer can be very narrow, essentially only signal is measured when the analyzer is tuned to each center band frequency   From these quantities the signal-to-noise (S/N) of each band can be calculated   From these S/N's, the weighting function (w_i) of each band is calculated by Equation 1, page 2-6, repeated below for convenience as Equation 7

$$w_i = 0 \quad \text{if } (S/N)_i < -12 \text{ dB}$$

$$(7)$$

$$w_i = \frac{(S/N)_i + 12}{30} \quad \text{if } -12 \text{ dB} \leq (S/N)_i \leq +18 \text{ dB}$$

$$w_i = 1 \quad \text{if } (S/N)_i > 18 \text{ dB}$$

SINE WAVE GENERATORS

SUMMING
AMPLIFIER

TAPE RECORDER
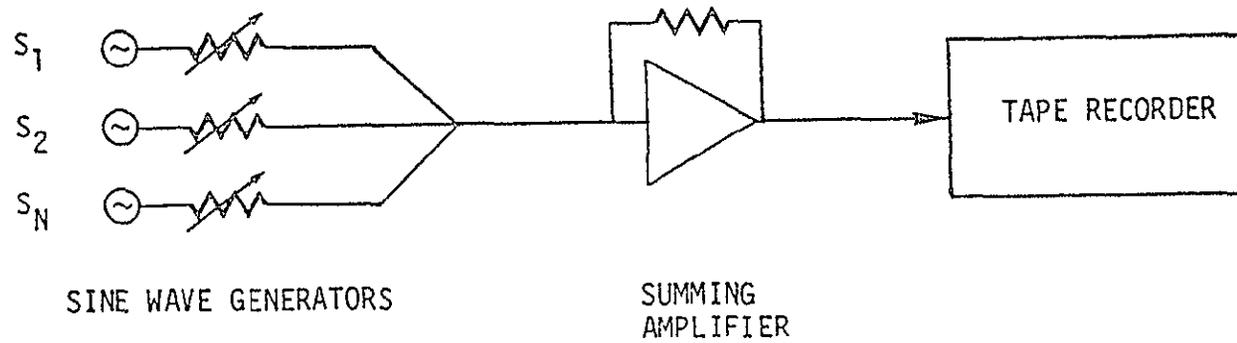
Figure 25   Discrete Input Frequency Method of Determining AI
System Input Mechanization
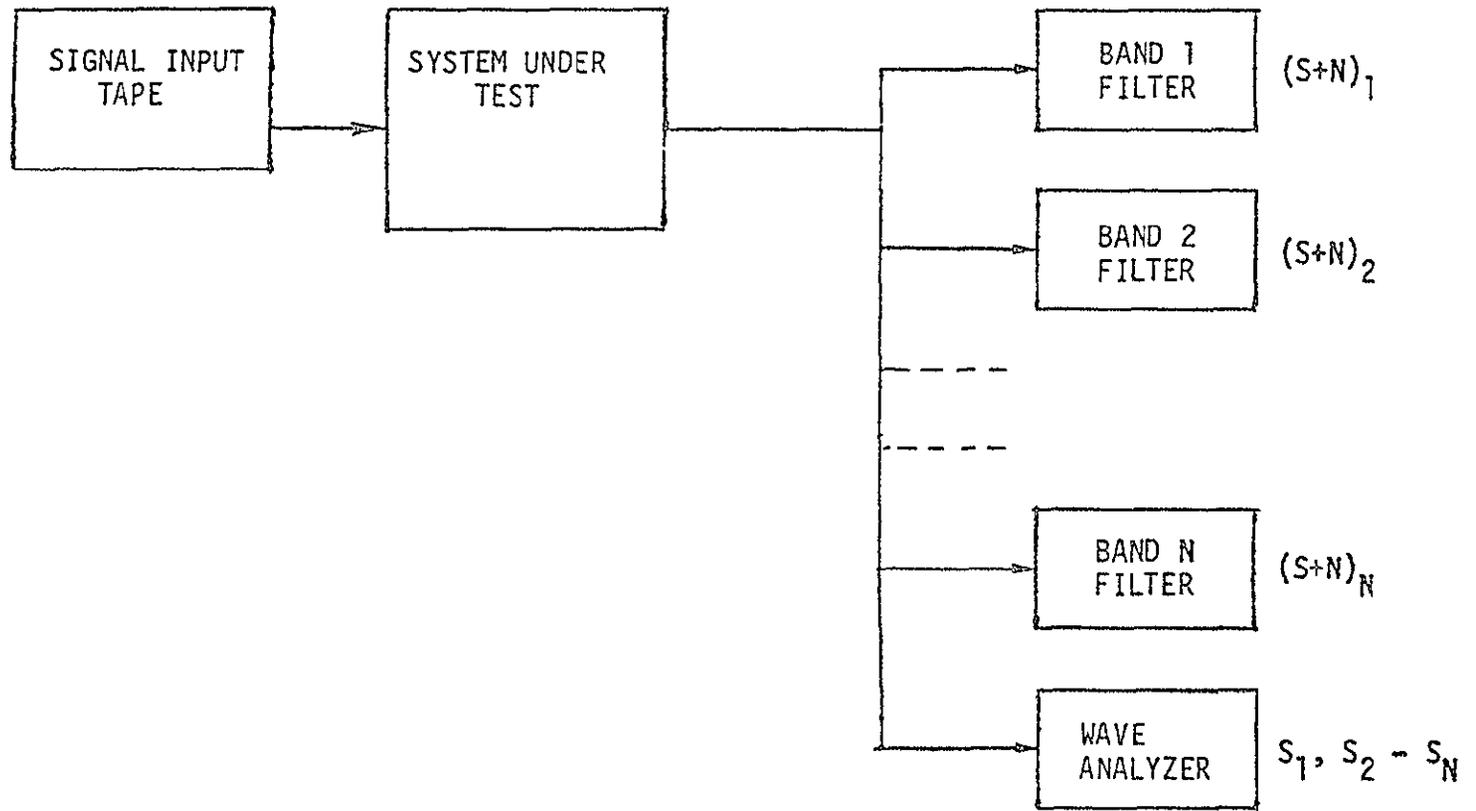
Figure 26   Simplified Method of Measuring AI

The articulation index (AI) if then computed from the weighting functions by Equation 2, page 2-6, repeated below as Equation 8

$$AI = \frac{1}{N} \sum_{i=1}^{N} W_i \tag{8}$$

where N = the number of equal importance articulation bands in the voice pass band of the system under test (N = 20 for the pass band of normal speech ranging from 200 Hz to 6100 Hz)

Figure 27 illustrates a more automated method of mechanizing the calculation of AI using a combination of analog and digital circuitry.

The output of the system under test is fed to a bank of filters which match the equal importance bands shown in Table 1 The output of each filter is fed to a solid state multiplexer, as well as to a very narrow band filter tuned to the center frequency of that band Thus, the output of the band filter represents signal plus noise (S+N), and the output of the narrow band filter is almost pure signal (S)

The multiplexer is illustrated in Figure 28, and is arranged to simultaneously sample the output S and S+N for each band in sequence, with each band being sampled for 05 seconds The total time for sequencing the entire voice band is 1 0 seconds as shown in Figure 29

After conversion into RMS voltages, the S_i and $(S+N)_i$ for each band are fed to a differential amplifier where the S_i is subtracted from the $(S+N)_i$ resulting in N_i for each band Both the S_i and N_i are fed to logarithmic amplifiers the output of which are representative of S_i in dB and N_i in dB, respectively The dB outputs of S_i and $(S+N)_i$ are applied to a differential amplifier resulting in $(S/N)_i$ in dB The output $(S/N)_i$ in dB is applied to a sample and hold circuit which is triggered by a delayed clock trigger The delay (with reference to the standard 20 Hz clock) is such that the voltage representing $(S/N)_i$ is sampled near the end of the 05 second gate time to ensure that the voltage has reached a stable value
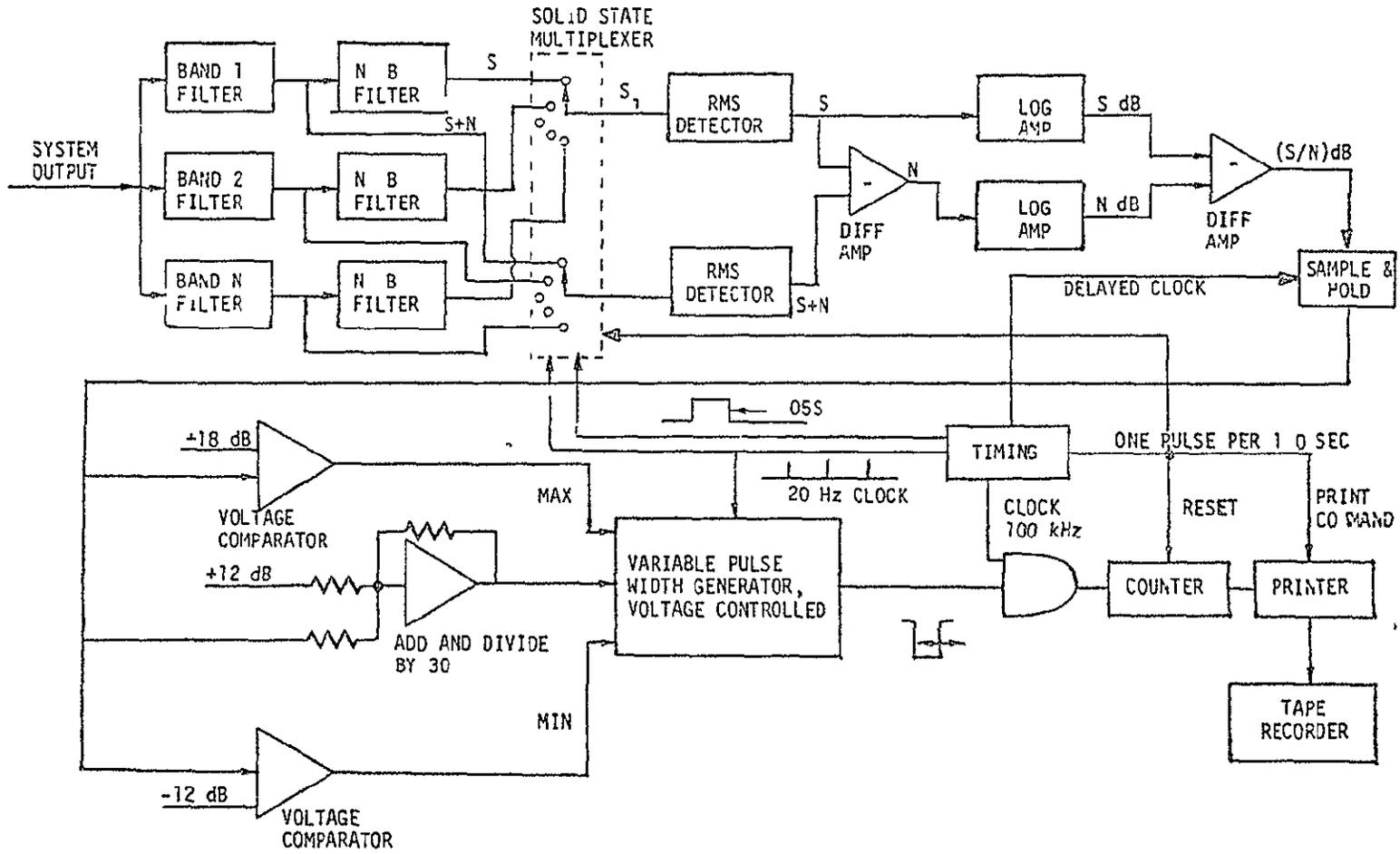
Figure 27   Mechanization of AI Computation
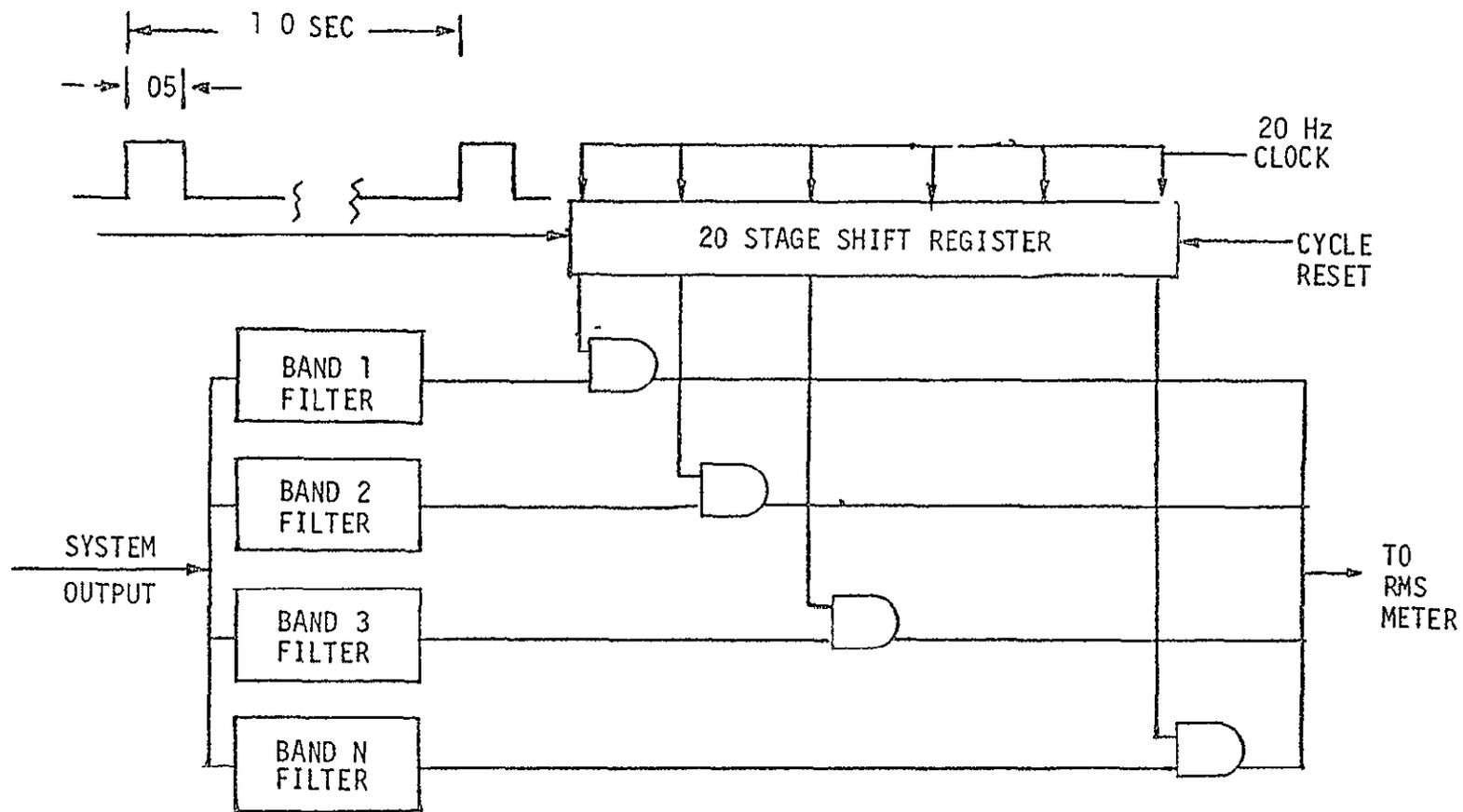
Figure 28   AI Mechanization Multiplexer Diagram
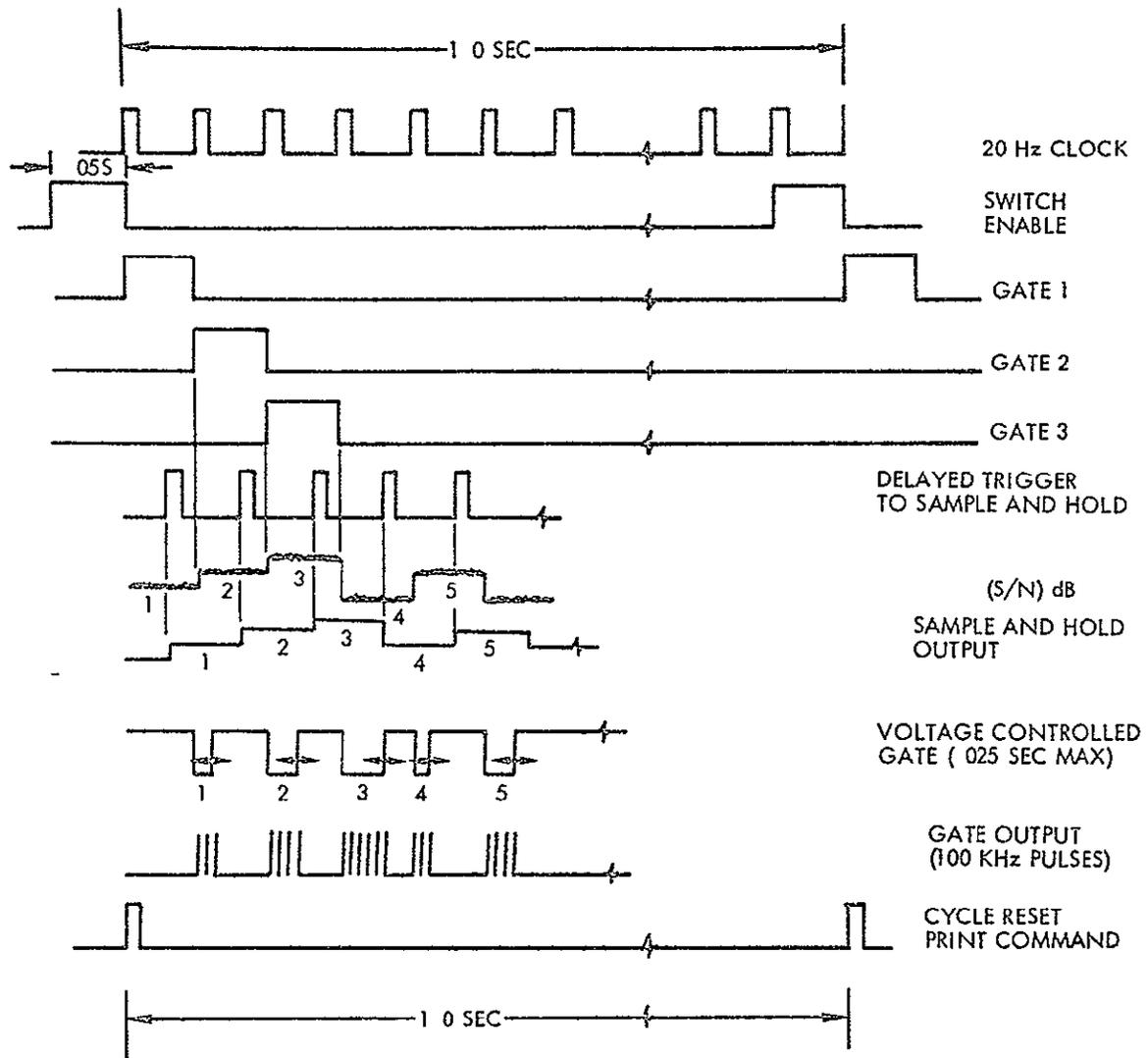
Figure 29   Timing Waveforms for AI Mechanization

The mechanization of Equation 7 is accomplished by applying the $(S/N)_1$ to three circuits in parallel, the outputs of which are fed to a variable pulse width generator which is voltage controlled   If the $(S/N)_1$ is less than -12 dB the variable width pulse generator remains biased in the off condition for the duration of that particular band input (05 seconds) If the $(S/N)_1$ is greater than -12 dB but less than +18 dB, the $(S/N)_1$ is added to +12 dB and divided by 30, and causes the output of the gate generator to be a proportional amount of the maximum width   If the (S/N) is equal to or greater than +18 dB, the variable width pulse generator is set to the maximum of 025 seconds for that band interval

The variable width pulse generator is triggered by the same 20 Hz clock that triggers the multiplexer   For each band, the width of the output pulse varies from 0 to 025 seconds, depending on the $(S/N)_1$ of that particular band   The maximum width of the variable width pulse generator is set to 025 seconds instead of the maximum allowable time of 05 seconds to prevent timing problems which could occur if successive $(S/N)_1$ outputs exceeded +18dB, equivalent to maximum width .

The output of the variable width pulse generator is used to gate 100 kHz pulses into a counter, where the count, proportional in each case to the W_1, is added until the cycle of 20 bands is completed, when it is reset (Figure 29)   The output of the counter over the entire cycle is also outputed on a printer resulting in permanent record of the AI

The frequency of the clock trigger acts as the scale factor (1/N) in Equation 2.   For the 20 band example, the 200 kHz used would result in a count of 100,000, equivalent to an AI of 1, if the $(S/N)_1$ in each band equaled or exceeded +18 dB

4 2.3 Alternate Mechanization

An alternate method of mechanizing the AI solution is shown in Figure 30   The method is the same as the preceeding example up to the determination of $(S/N)_1$ Here the values of $(S/N)_1$ between -12 dB and +18 dB would be converted into digital form after mechanization of formula 2, and stored in shift registers, one for each band   The values of $(S/N)_1$ greater than

SYSTEM OUTPUT

BAND 1 FILTER → NARROW BAND FILTER

BAND 2 FILTER → NARROW BAND FILTER

BAND N FILTER → NARROW BAND FILTER

MULTIPLEXER

RMS V M → LOG AMP → S dB

RMS V M → LOG AMP → (S+N)dB

DIFF AMP → (S/N)dB

TIMING

+18 dB → DIFF AMP → SET ONE IF GREATER THAN REF

+12 dB → ADD → − 30 → A/D CONV

−12 dB → DIFF AMP → SET ZERO IF LESS THAN REF

T
S → BAND 1 SHIFT REGISTER
R

T
S → BAND 2 SHIFT REGISTER
R

T
S → BAND N SHIFT REGISTER
R

ADD → ÷ N → AI
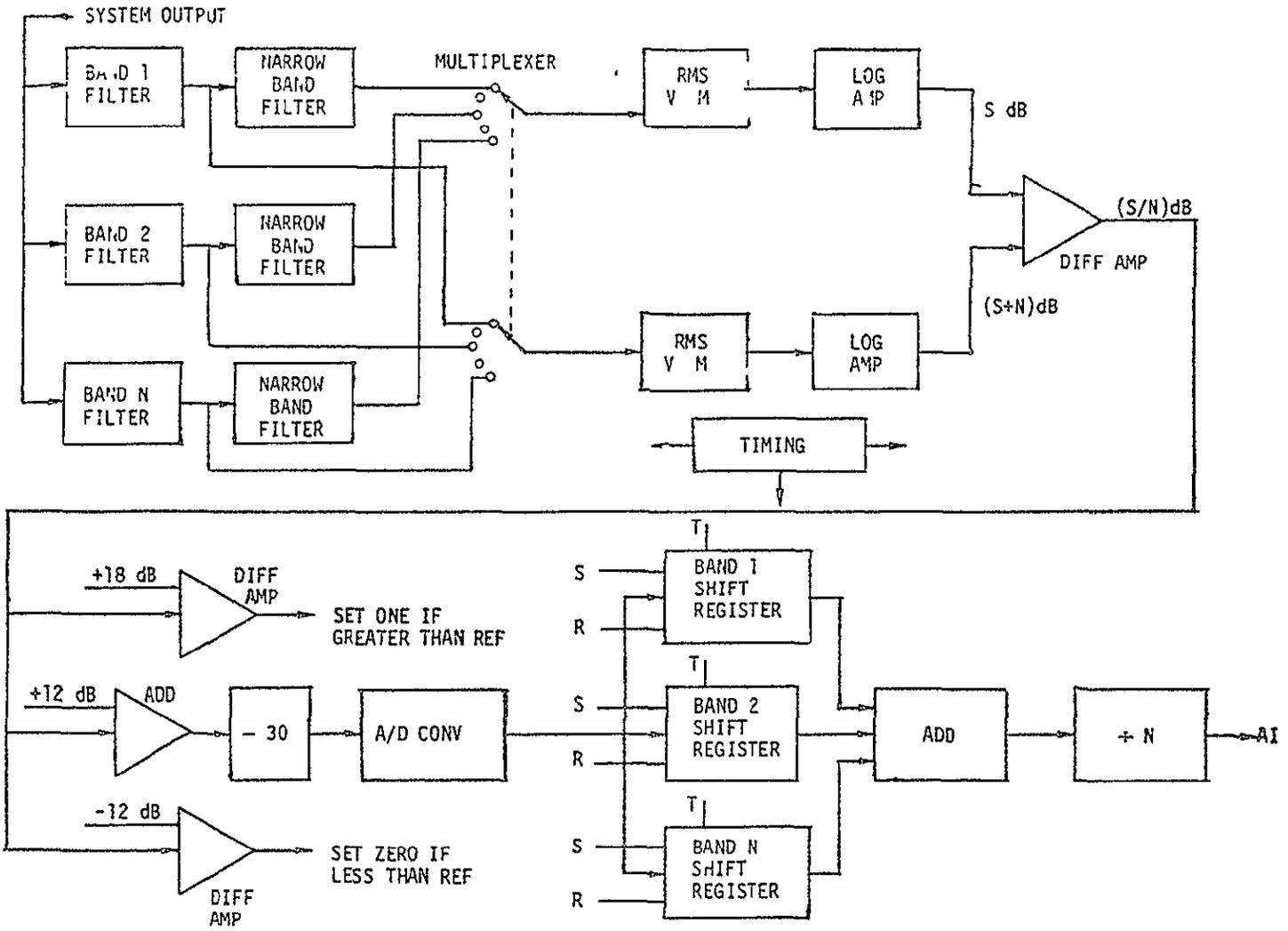
Figure 30 Alternate Method of AI Mechanization

+18 dB would set all 1's in the register, and values less than 12 dB would set all 0's. At the end of the N band cycle the outputs of the N registers would be added and the sum divided by N to provide the AI.

4 3 VIDEO SIGNAL TO NOISE RATIO MEASUREMENT USING WEIGHTED NOISE

4 3 1 Input Signals

The picture signal-to-noise ratio (SNR) is defined by

$$ SNR = \left(\frac{sync\ to\ white\ level\ video\ voltage}{weighted\ RMS\ video\ noise\ voltage} \right)^2 $$

(Reference 9)   Since this definition sets as limits the allowable signal voltage swing peak-to-peak, the input video signal level must be determined by the system under test to produce the maximum demodulated signal   The actual limits will be determined by the type of modulation, permissable carrier deviations, etc , of the system under test   The picture SNR defined above is based on the use of actual video input (selected still pictures scanned by a slide viewer)   However, the picture SNR definition itself is not picture dependent, since it is based on a maximum allowable signal level   When use is made of simulation for actual video input, this fact must be taken into account   The picture SNR criteria when used with a simulated video input, must be used in combination with other criteria, including video bandwidth   A standard bandwidth criteria is an essentially flat ($\pm$ 25 dB) response from d c to 4 MHz (Reference 13)

The composite video signal input is composed of actual video signals, blanking and synchronizing signals (Figure 31)   In order to simulate the effects of the synchronizing signals and high frequency video frequencies, a square wave input of 15 75 kHz (representative of the television line frequency) is proposed as test signal input

4 3.2 Output Signal Mechanization

One method mechanizing the measurement of picture SNR is shown in Figure 32   The input signal square wave, properly attenuated, is fed to the system under test through a solid state switch S_1, operating synchronously with switch S_2 which switches the output of the system under test
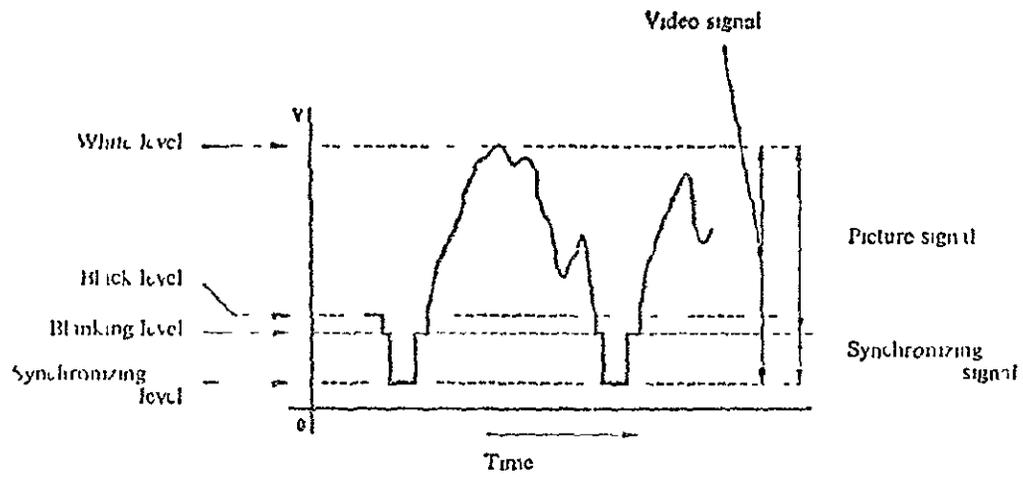
Nominal Value of Voltage from White to
Sync Level = 1 Volt

Figure 31. Typical Video Signal
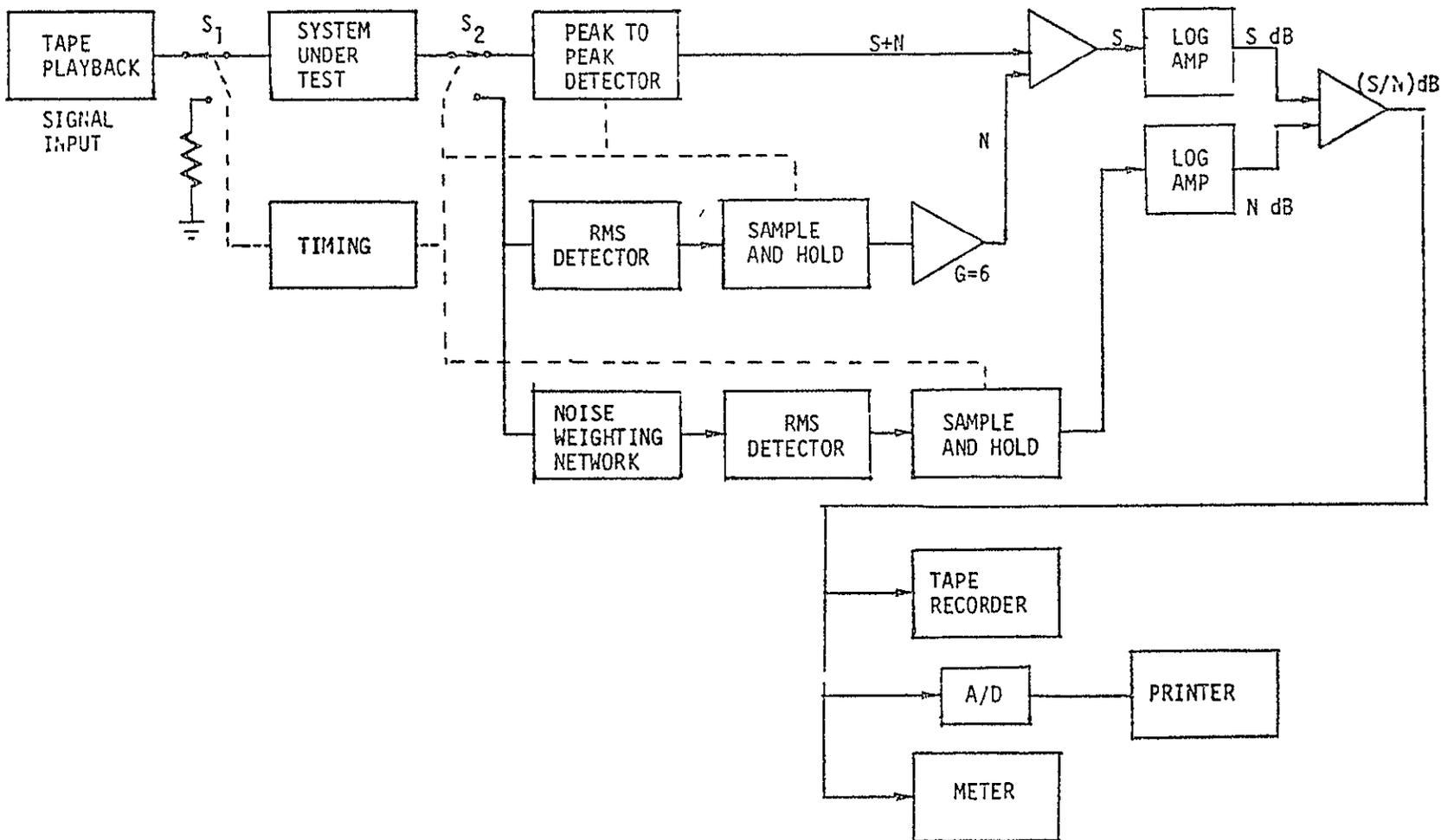
Figure 32   Mechanization of Video Picture
SNR Using Noise Weighting

to the test equipment. In the switch position shown in Figure 32, the input signal S is fed into the system under test, and the output signal plus noise (S+N) is fed to a peak-to-peak detector. The output of the peak-to-peak detector is fed to a differential amplifier. The solid state switches S_1 and S_2 are controlled by the timing circuitry, and are opened and closed during alternate one half second periods. The timing circuits also provide the sampling pulse to the sample and hold and peak-to-peak detector circuits. Figure 33 shows a means of providing the switching action for S_1 and S_2 and a timing diagram showing the relative timing of the S and S+N signals and the sampling pulses.

In the other position, switch S_1 terminates the input to the system under test in the correct impedance and S_2 allows the system noise to be fed to an RMS detector. The output of the RMS detector, system noise (N), is sampled during this period and held for comparison with the S+N. Since the amplitude of the peak value of white Gaussian noise is approximately three times its RMS value, the noise from the sample and hold circuit is multiplied by a factor of six to enable its comparison with the peak-to-peak value of the S+N.

The output of the differential amplifier is the difference between (S+N) and N, and thus represents the signal S. This output S is fed through a logarithmic amplifier providing an output signal S in dB. The use of this circuit to subtract the noise from the signal pulse noise is a refinement which may not be necessary when the values of signal to noise ratios which are likely to be encountered are considered. A representative value of SNR required to produce a "good" picture is generally taken to be approximately 50 dB for commerical television, while the lowest value to be tolerated is somewhere in the region of 30 dB (Reference 3). In other words the noise to be expected is about one percent of the signal, or lower. Thus, the error involved in considering (S+N) to be S is small.

The noise output from switch S_2 is fed to a noise weighting network which has a frequency response shown in Figure 8. The weighted noise is fed to a RMS detector and a sample and hold circuit and is converted into
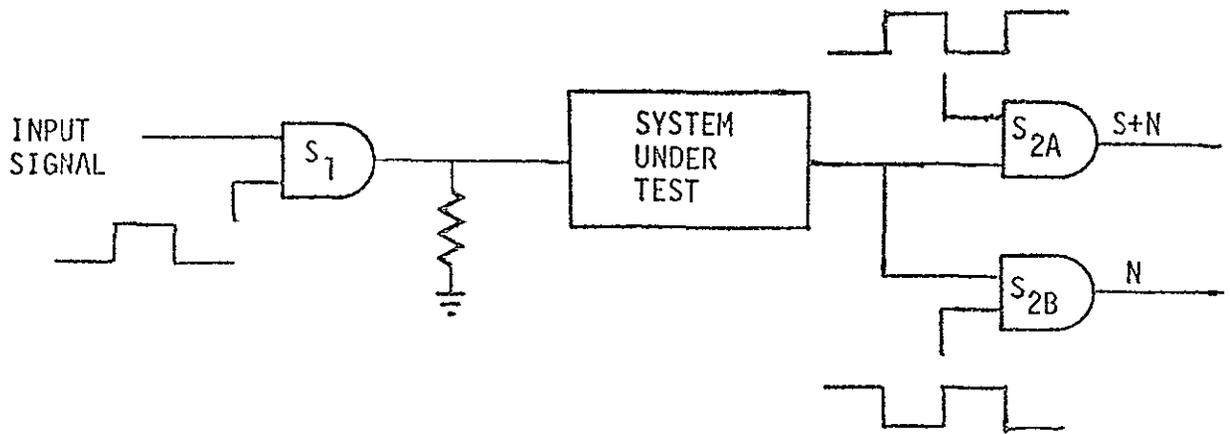
SOLID STATE SWITCHES
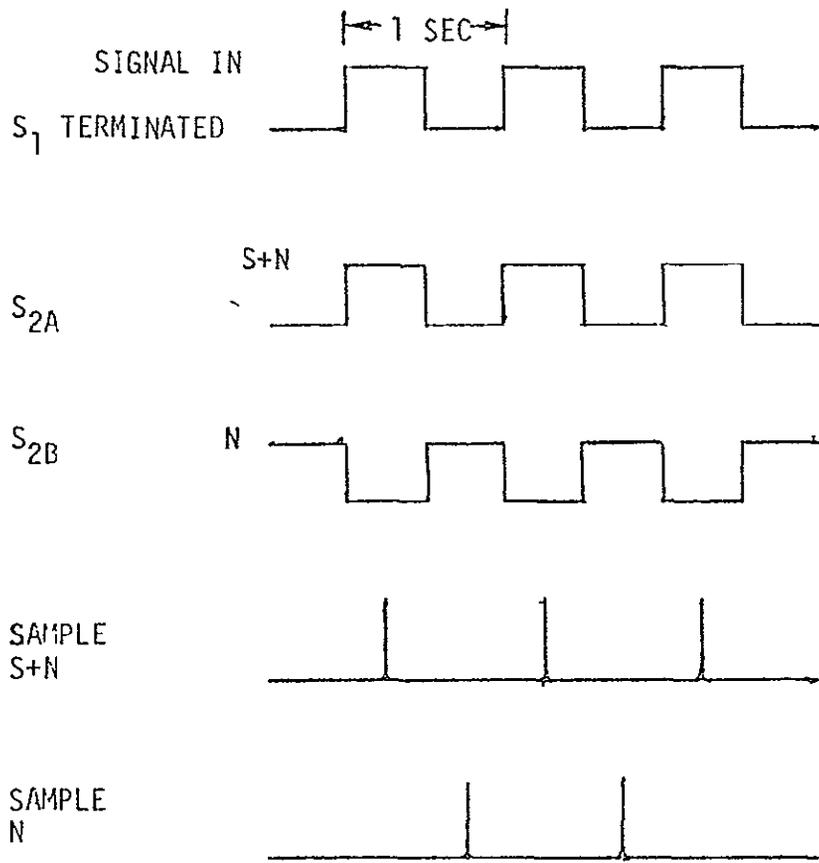
TIMING DIAGRAM

Figure 33   Video Picture SNR Switching Diagrams

dB by a logarithmic amplifier and fed to a differential amplifier where it is compared to the peak-to-peak signa, S, to produce the weighted picture SNR.

The output weighted picture SNR = $\frac{S_p}{N_{rms}}$ from the differential amplifier is applied to the tape recorder either in analog form or digital form. In addition, it can be displayed on an analog or digital voltmeter.

4 3.3 Calibration Signals

Aside from the measurement of the picture SNR, the principal parameter to be determined is video frequency response (Reference 3 shows the required response to be flat to 4 MHz within $\pm$ 0 25 dB for commerical television) Figure 34 shows a calibration circuit for measuring video frequency response. The calibration signals will include a swept frequency of constant amplitude. The output will be displayed on an oscilloscope whose time base is set by the sweep voltage controlling the frequency The output frequency response will also be recorded

Also shown in Figure 34 is a means of mechanizing the output to provide an indication that the response has exceeded the $\pm$ 0 25 dB limits of flatness. The output from the system is passed through a 100 kHz filter to be used as a reference. After detection by an RMS detector the 100 kHz response is sampled and held A portion of the 100 kHz reference proportional to 0.25 dB is also obtained and fed to the final comparator stage. The unfiltered output is detected and fed to a window comparator circuit where it is compared with the output of the 100 kHz reference The input stage of the window comparator is an obsolute value circuit where the absolute value of the difference of between the reference and unfiltered output is obtained. This is compared to the 0 25 dB reference in a differential amplifier. When the absolute value of the differential exceeds 0 25 dB, an indication is obtained which can be used to produce a visual or other display

VOLTAGE
SOURCE

ABSOLUTE
VALUE CIRCUIT

100 kHz
FILTER

RMS
DET

SAMPLE
AND HOLD

INDICATOR

OUTPUT

FROM
SYSTEM
WITH
SWEPT
FREQ
INPUT

RMS
DET

WINDOW
COMPARATOR

TAPE
RECORDER

CALIBRATION CIRCUIT

OSCILLOSCOPE
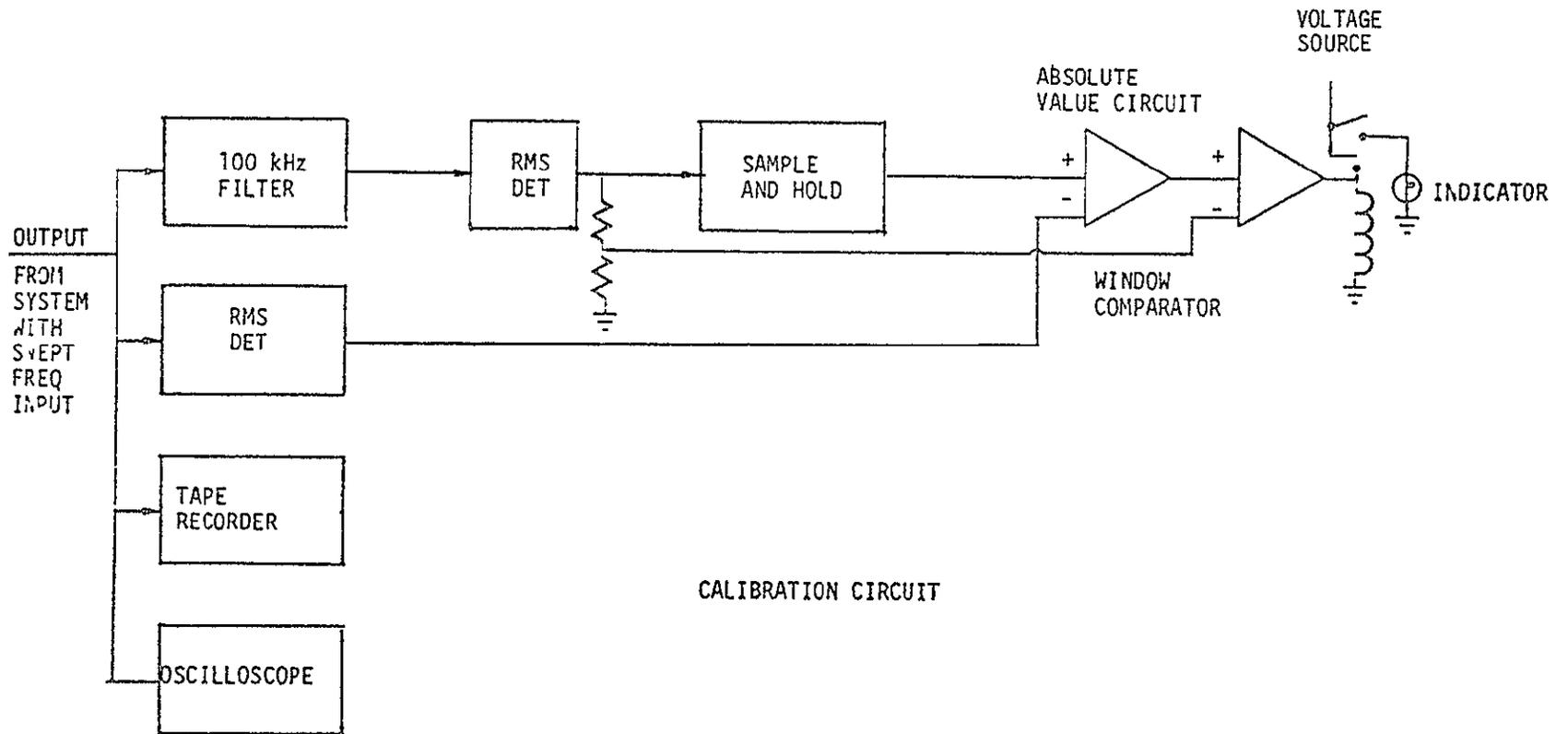
Figure 34   Calibration Circuit for Video SNR

4 4   ESTIMATION OF BIT ERROR RATE BY EXTRAPOLATION OF PSEUDO ERROR RATE

4.4 1   General Discussion

In digital data systems, the Bit Error Rate (BER) is the performance criterion most generally specified   BER is most simply measured by comparing the output to a true copy of the input, and counting the errors over a period of time   In a system with a very small actual error rate, the time required to count a statistically significant number of errors may become unacceptably large   One method of reducing this time is based on the generation of psuedo error rates which are greater than the actual rate, and the estimation of the actual error rate by extrapolation

The time required to measure error rates in a digital data system can be considerably reduced by means of an external variable threshold decision device connected in parallel with the decision device in the data system under test   By biasing the decision threshold of the external decision device, its error rate can be made to greatly exceed that of the unbiased decision device in the system under test   By plotting the logarithm of this pseudo-error rate from the external decision device versus the threshold bias, and extrapolating the resulting curve to the error rate for zero threshold bias, an estimate of the true error rate of the system under test can be obtained

The measurement of the average bit error rate by the pseudo-error rate extrapolation method is most simply implemented by a manual data acquisition system of the type shown in Figure 35   In this manual system, a threshold parameter, K, is defined by

$$K = \frac{M}{A} ,$$ where A is the peak signal amplitude at the external decision device, and M is the threshold level of the decision device

Thus, K is the fractional threshold setting   Values of error rate, P_T, are measured at each of several values of K   The resulting data is plotted
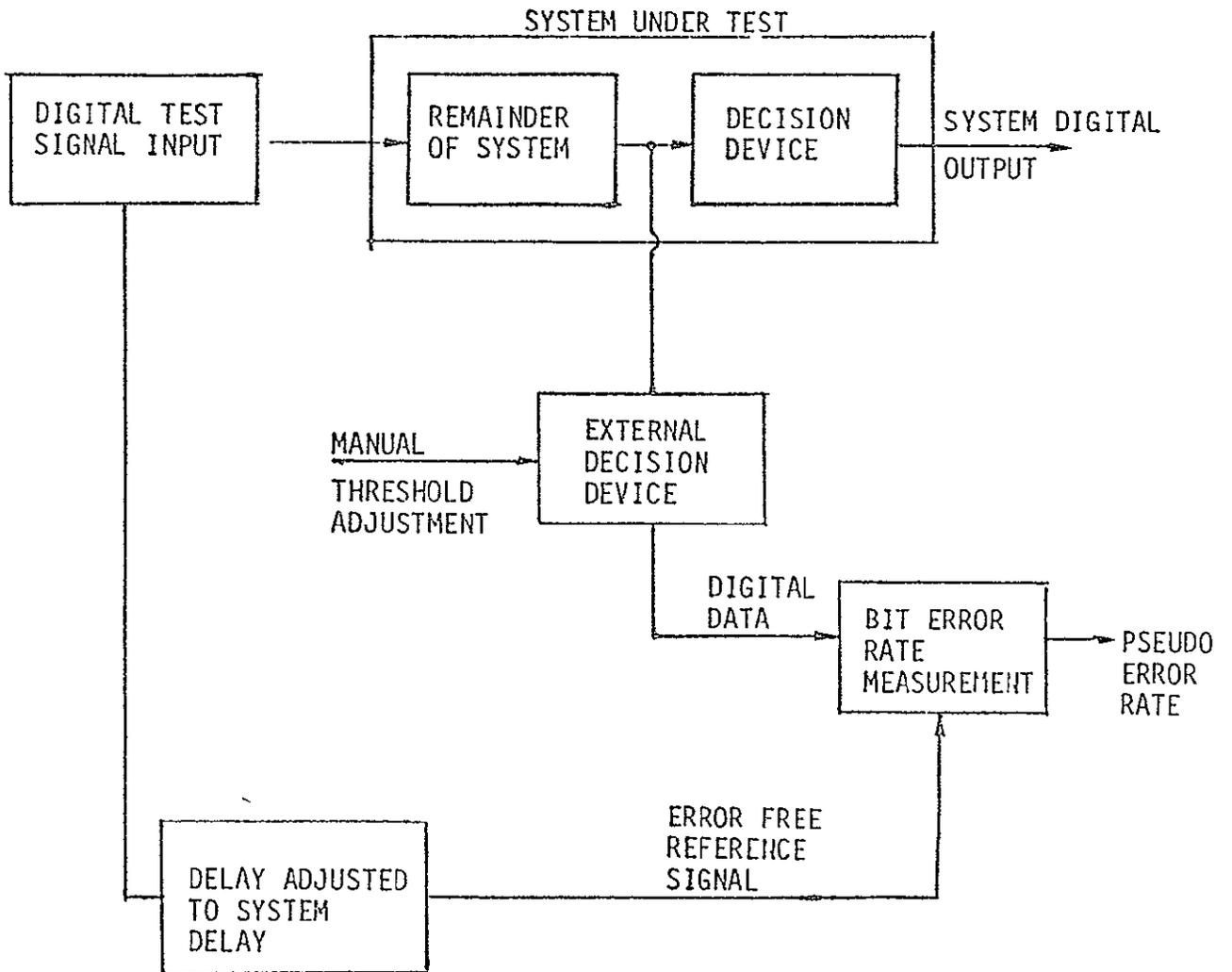
Figure 35. Generation of Pseudo Error Rate

as K versus $\log_{10} P_T$ and the resulting curve extrapolated to K=0 The P_T corresponding to K=0 is an estimate of the true error rate.

A clearer picture of the operation of this technique can be gained by considering the example of a binary PSK demodulator operating in a Gaussian noise environment If the threshold is modified so as to increase the number of 1's errors, then the probability of 1's error can be determined by computing the cross-hatched area in the sketch of Figure 36(a) This probability is given by

$$P_{E,1} = \frac{1}{\sqrt{2\pi}\,\sigma} \int_{-\infty}^{M} e^{-\frac{(X-A)^2}{2\sigma^2}} \, d_X$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{A}{\sigma}(K-1)} e^{-u^2/2} \, d\mu \tag{9}$$

Similarly, the probability of a 0's error is given by the cross-hatched area in Figure 36(b) This probability is given by

$$P_{E,0} = \frac{1}{\sqrt{2\pi}\,\sigma} \int_{M}^{\infty} e^{-\frac{(X+A)^2}{2\sigma^2}} \, d_X$$

$$= \frac{1}{\sqrt{2\pi}} \int_{\frac{A}{\sigma}(K+1)}^{\infty} e^{-\mu^2/2} \, d\mu \tag{10}$$

where σ is the rms value of the noise and
M, A, and K are as previously defined

The total probability of error, P_T, is

$$P_T = \frac{1}{2}(P_{E,1} + P_{E,0}) \tag{11}$$

4-22
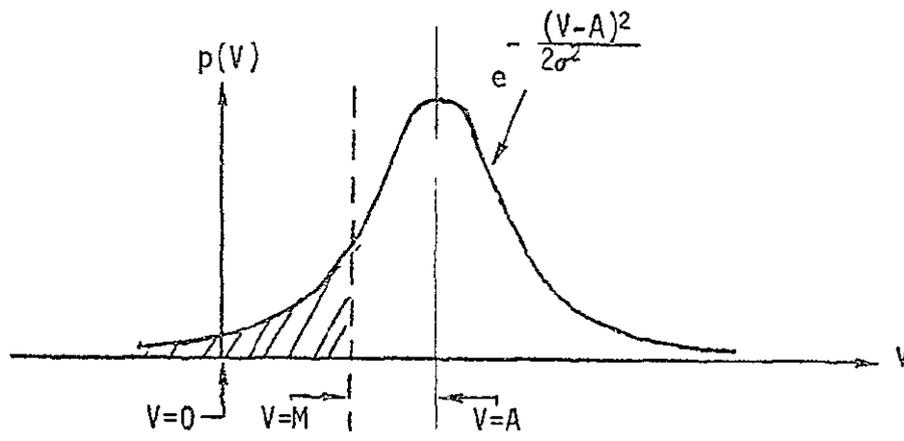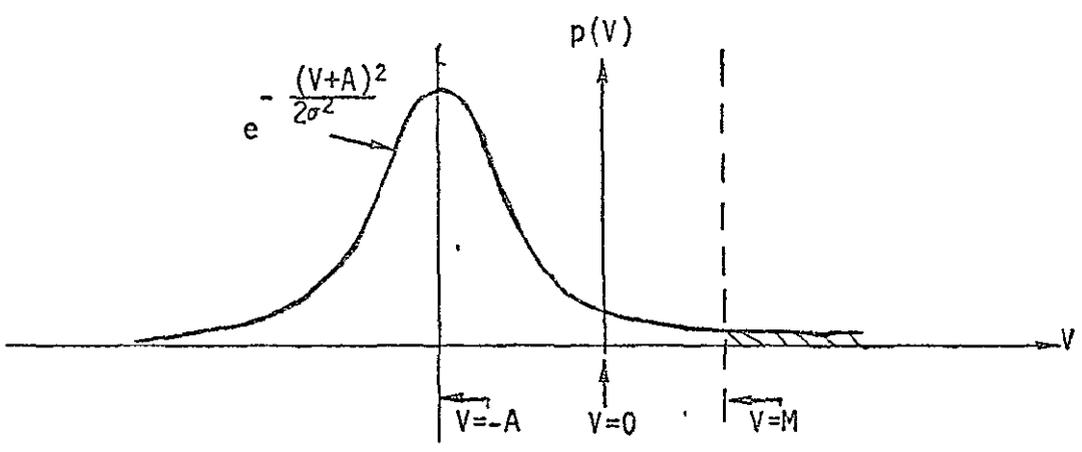
Figure 36   Error Probability curves for External Threshold
           Device Biased to Increase the 1's Error Rate

Table 6 gives a tabulation of $\log_{10} P_T$ as a function of the parameter, K. The values were obtained by evaluating equations (9) and (10) for a $A/\sigma = 4\,0$ This data is plotted in Figure 37 which shows the essentially straight line relationship that exists between K and $\log_{10} P_{E,T}$. The extrapolated curve gives a value of $\log_{10} P_{E,T}$ of $-4\,49$ which corresponds closely to the calculated value of $-4\,50$ given in Table 6

K	$\left|\frac{A}{\sigma}(K-1)\right|$	$P_{E,1}$	$\left|\frac{A}{\sigma}(K+1)\right|$	$P_{E,0}$	$P_{E,T}$	$\log_{10} P_{E,T}$
0	4	$3\,2 \times 10^{-5}$	4	$3\,2 \times 10^{-5}$	$3\,2 \times 10^{-5}$	$-4\,50$
0 2	3 2	$6\,9 \times 10^{-4}$	4 8	$7\,9 \times 10^{-7}$	$3\,5 \times 10^{-4}$	$-3\,46$
0 3	2 8	$2\,6 \times 10^{-3}$	5 2	1×10^{-7}	$1\,3 \times 10^{-3}$	$-2\,89$
0 4	2 4	$8\,2 \times 10^{-3}$	5 6	$1\,1 \times 10^{-8}$	$4\,1 \times 10^{-3}$	$-2\,39$

Table 6   Tabulation of Error Probabilities for Signal with Peak signal to rms noise $(A/\sigma) = 4$

4 4 2   System Mechanization

The pseudo-error rate extrapolation technique can be readily automated to provide a device whose output is an estimate of the true error rate of the system under test. The general arrangement is shown in the block diagram of Figure 38

In this arrangement, the three separate decision devices have three different values of K set for the duration of the test The error counters compare the outputs of the decision devices with a true copy of the original data and supply a number proportional to pseudo-error, $P_{E,K}$, to the extrapolator The extrapolator forms the logarithm of each $P_{E,K}$ value and forms a least squares straight line fit to the three values of $\log(P_{E,K})$. The intercept of this line with the $\log(P_E)$ axis is determined, the antilogarithm formed and supplied as an output estimate of the true error rate
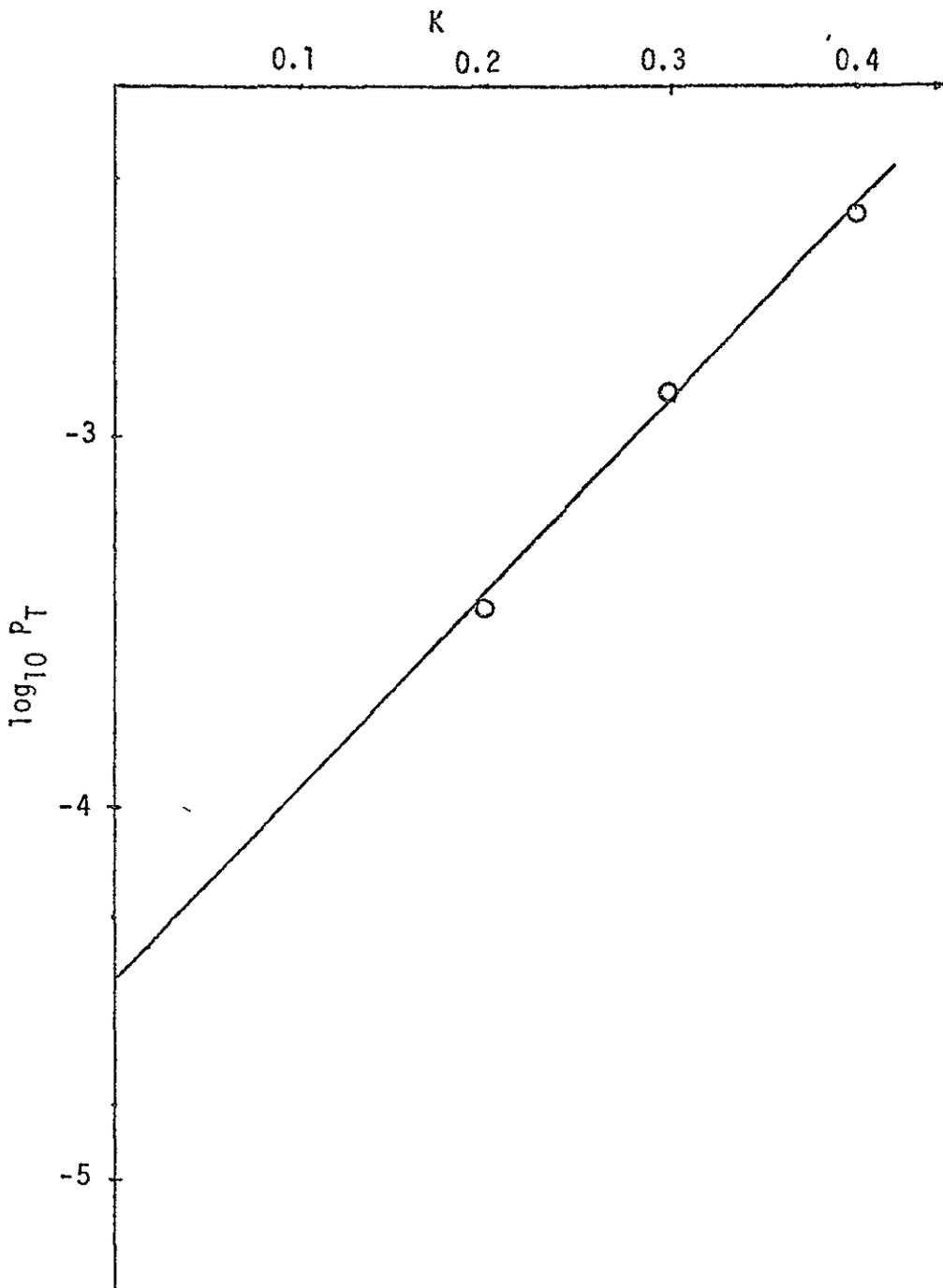
Figure 37   Logarithm of Estimated Error Rate, P_T, versus K, with Peak Signal to RMS Noise = 4
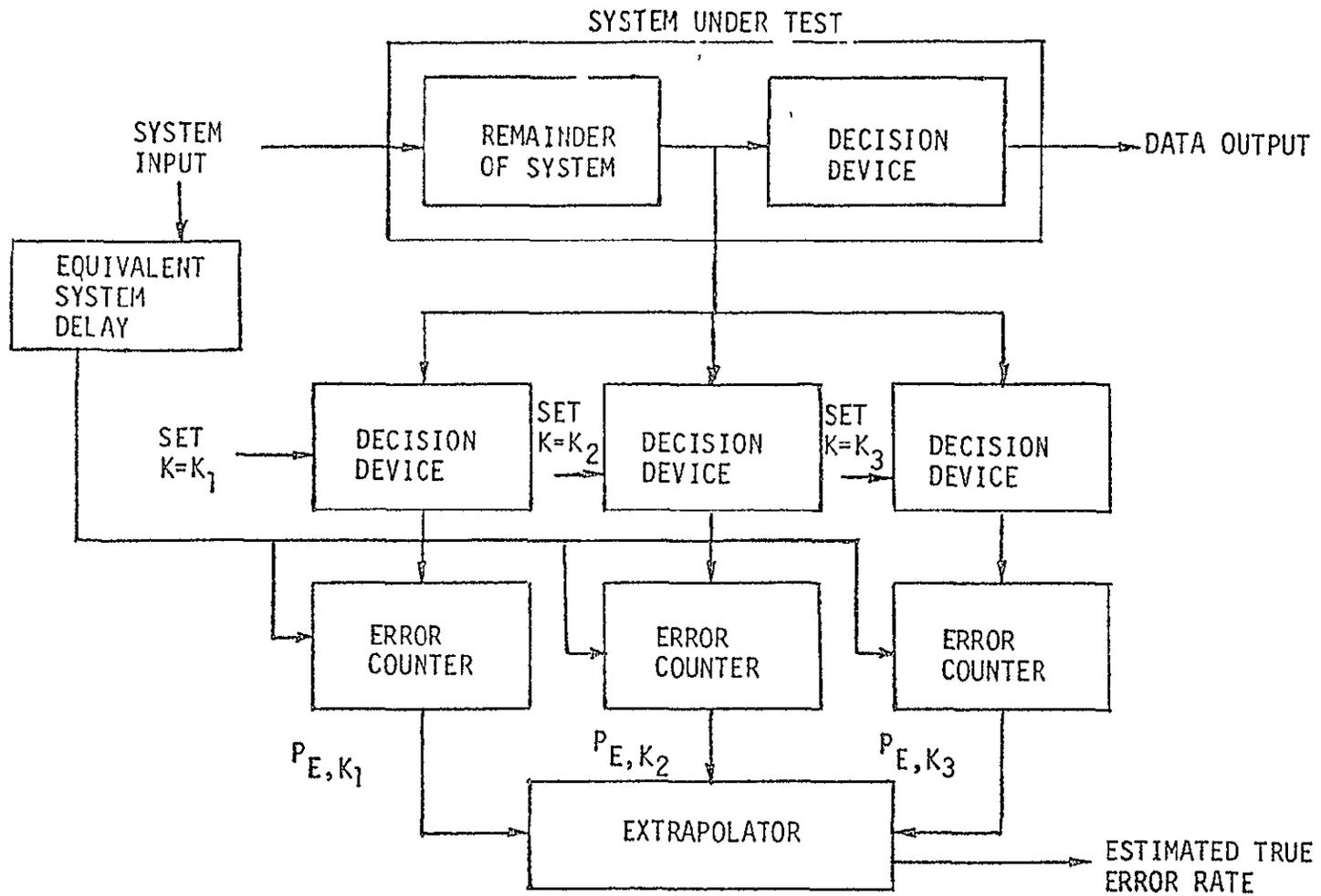
Figure 38   Error Rate Estimator

If the equation of the least squares fit straight line is

$$\log P_C = a_o + a_1 K_1 \tag{12}$$

then the desired intercept is a_o The value of a_o can be computed from

$$a_o = \frac{\left(\sum_{i=1}^{3} P_{E,K_1}\right)\left(\sum_{i=1}^{3} K_1^2\right) - \left(\sum_{i=1}^{3} K_1\right)\left(\sum_{i=1}^{3} P_{E,K_1} K_1\right)}{3 \sum_{i=1}^{3} K_1^2 - \left(\sum_{i=1}^{3} K_1\right)^2} \tag{13}$$

(Reference 17)

An analog mechanization of this solution for a_o is shown in Figure 39.

In Figure 39, the quantities B, C, D, R are convenient scaling factors which allow convenient values to be assigned to the various summing and feedback resistors
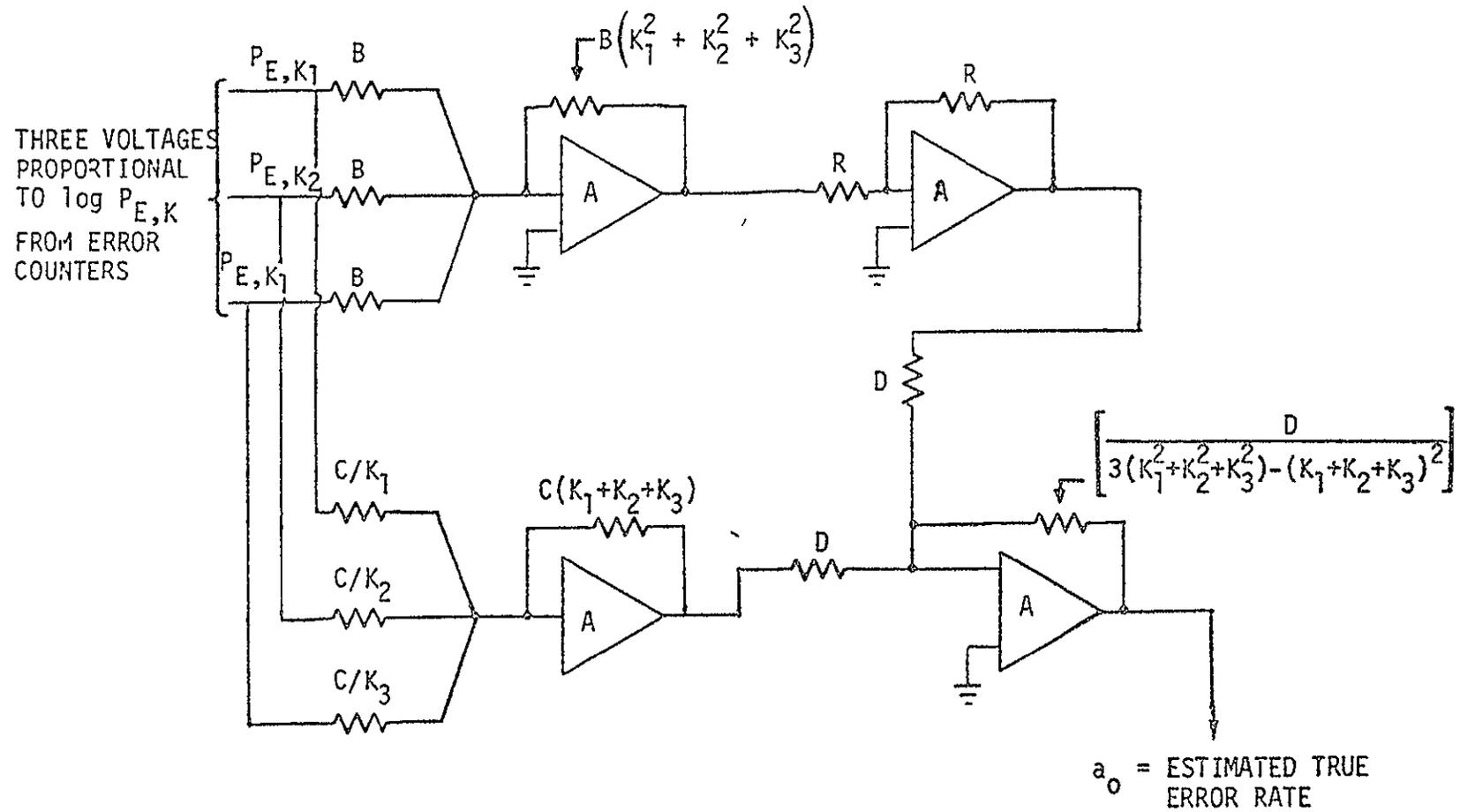
Figure 39. Pseudo Error Extrapolator Mechanization

5. CONCLUSIONS AND RECOMMENDATIONS

The following conclusions are drawn as results of the study described in this report

1. In the area of voice communication, the use of the articulation index, which represents a weighted average of the SNR's of a number of frequency bands comprising the voice band, emerges as a clear choice of performance criteria.

   Although the AI of a particular system must be related to word intelligibility by empirical means, the AI itself as a criteria is one which is repeatable, reasonably easy to measure, and is more definitive than a gross SNR taken over the entire voice band

2. As regards video information, the choice is not so clear Picture quality evaluation is more subjective and is not as easily defined as word intelligibility The use of weighted picture SNR appears to be the best answer as a repeatable, easily measured criteria Weighted picture SNR does appear to have a reasonably linear relationship to picture quality, as determined by several investigators

3. For digital data, average bit error rate (BER) remains the obvious standard criteria for uncoded systems where the distribution of errors is not important

An evaluation of a number of methods of mechanizing criteria evaluation results in the following

1. For voice systems, a system of mechanizing the measurement of AI by determining the response of the system under test to a series of discrete frequencies centered in the "equal importance bands" is proposed as offering a reasonable method of achieving the desired result.

2. Use of a square wave with frequency equal to the standard television line frequency and use of a noise weighting filter are the chief features of the method suggested for mechanizing picture SNR measurements.

3   A method of mechanizing the use of psuedo error rates to reduce the time required to determine actual bit error rates in digital data systems with low BER's is proposed   Pseudo error rates which are large with respect to the actual error rates are created by biasing the "mark" and "space" decision threshold detectors away from the average reference. The actual error rate is determined from the pseudo rate

APPENDIX A

COST ESTIMATES

In order to provide a gross estimate of what might be involved in
realizing the hardware mechanization of the techniques described in Section
4, the following parts lists and cost estimates are presented   They are
intended as guidelines and are furnished for information only   Based on
the use of commercially available test equipment and off-the-shelf packaged
units where possible, they do not include preparation of the test signal
tapes or the tape recorder/playback system   No labor cost estimates are
included

Voice Articulation Index Mechanization

20 "Equal Importance Band" Filters @ $125		$ 2,500 00
20 Narrow Band Filters @ $125		2,500.00
Multiplexer (shift register and gates)		250 00
2 RMS Meters @ $575		1,150 00
2 Logarithmic Amplifiers @ $420		840 00
10 Operational Amplifiers @ $50		500 00
Timing Circuitry		100 00
Counter		1,000 00
Printer		1,200 00
Rack, Power Supply, Cabling, etc		500 00
	Total	$10,540 00

VIDEO SNR MECHANIZATION

	Peak-to-Peak Detector	$ 150 00
2	RMS Meters at $575	1,150 00
8	Operational Amplifiers @ $50	400 00
2	Logarithmic Amplifiers @ $420	840.00
	Timing and Switching	150 00
	Noise Weighting Network	50.00
	Analog to Digital Converter	500 00
	Pack, Power Supply, Cabling, etc	300 00
	Total	$ 3,454 00

VIDEO SNR CALIBRATION CIRCUIT

	100 kHz Filter	$ 150 00
?	RMS Meters @ $575 00	1,150 00
4	Operational Amplifier @ $50	150 00
	Indication Circuits	10.00
	Rack, Power Supply, Cabling, etc	250 00
	Total	$ 1,710.00

BER ESTIMATION MECHANIZATION

3	Decision Circuits @ $150	$ 450 00
3	Error Circuits @ $500	1,500 00
	Extrapolation Circuits	250 00
	Rack, Power Supply, Cabling, etc	250 00
	Total	$ 2,450 00

REFERENCES

1. Project Technical Report, "Survey of Performance Criteria for Voice, Television, and Digital Data Transmission Systems," TRW Document No. 17618-H123-R0-00, 30 March 1971.

2. Project Technical Report, "New Criteria Development," TRW Document No. 17618-H174-R0-00, 29 June 1971.

3. EIA Standard RS-250-A, "Electrical Performance Standards for Television Relay Facilities," Electronic Industries Association, February, 1967

4. USA Standard, "Method for Measurement of Monosyllabic Word Intelligibility", May 25, 1960.

5 Harvey Fletcher, "Speech and Hearing in Communication," D Van Nostrand Company, Inc., 1935.

6. R D Prosser, et. al., "Quality Grading of Impaired Television Pictures," Proc. IEE, Vol. III, No 3, March 1964

7. J M Barstow, H N Christopher, "The Measurement of Random Video Interference to Monochrome and Color Television Pictures," Proc. AIEE, Pages 313-320, November 1962.

8. A. D. Fowler, "Observer Reaction to Low-Frequency Interference in Television Pictures," Proc IRE, October 1951.

9. J. Jansen, P. L. Jordan, et al., "Television Broadcast Satellite Study," Section 2, TRW Report No 08848-6002-R0-00, October 24, 1969

10. William R Bennett and James R. Davey, "Data Transmission," McGraw-Hill Book Company, 1965.

11. D J Gooding, "Performance Monitor Techniques for Digital Receivers Based on Extropolation of Error Rate," IEEE Transactions on Communication Technology, Vol. COM-16, No. 3, June 1968.

12 Charles A Rohde, "Estimation of Bit Probability of Error Using Sync Word Error Rate Data," IEEE Transactions on Communication Technology, December 1970.

13 Technical Report, "Development of a Speech-to-Noise Ratio Measurement Utilizing Analog Techniques," Philco-Ford Document PHO-TN248, 6 September 1968.

14. Technical Report "Development of a Speech-to-Noise Ratio Measurement Utilizing Digital Techniques," Philco-Ford Document PHO-TN228, 12 June 1968.

15. Draft Report "Technical Characteristics of Systems Providing Communication and/or Radio Determination Using Satellite Techniques for Aircraft and/or Ships," Document No. USSG IV/W-1142, 1 August 1970

16 J. O Weston, "Transmission of Television by Pulse Code Modulation," Electrical Communication, Vol. 42, No 2, 1967.

17. Murray R. Spiegel, "Statistics," Schaum Publishing Company, 1961.

18. Leo L. Beranek, "The Design of Speech Communications Systems," Proceedings of the I.R E., Vol. 35, 1947, pp. 880-890.

19. Reference Data for Engineers, 4th Edition, International Telephone and Telegraph Corporation, pp. 872-878

20 R L Craiglow, "A Study of the Effects of Elementary Processing Techniques on the Intelligibility of Speech in Noise," Section II, Collins Radio Research Report CRR 201, Reprinted January 1965

21 N R French and J C Steinberg, "Factors Governing the Intelligibility of Speech Sounds," Journal of the Acoustical Society of America, Vol 19, No 1, January 1947

22 Gorden L Fredendall, W. L Behrend, "Picture Quality - Procedures for Evaluating Subjective Effects of Interference," Proc IRE, June 1960

23 L. E. Weaver, "The Quality Rating of Color Television Pictures," Journal of the SMPTE, Volume 77, Pages 610-613, June 1968

24 D E Pearson, "A Realistic Model for Visual Communication Systems," Proc IECE, Volume 55, No. 3, Pages 380-389, March 1967

25. M. H Kubba, "Automatic Picture Detail Detection in Presence of Random Noise," Proc. IEEE, Pages 1518-1523, November 1963

26 R C Brainard, F W. Mounts, B. Prasada, "Low-Resolution TV Subjective Effects of Frame Repetition of Picture Replenishment", The Bell System Technical Journal, Pages 261-271, January 1967

27 F W Mounts and D E Pearson, "Apparent Increase in Noise Level When Television Pictures are Frame-Repeated," The Bell System Technical Journal, Pages 527-539, March 1969.

28 Peter T Roth, "Effective Measurements Using Digital Signal Analysis," IEEE Spectrum, April 1971, pp 62-70

29 NASA MSC Report "Electronic Systems Test Program, LM Phase III Television Report," MSC-EE7-70-207, 17 April 1970.